**Development of genetic tools for large-scale deletions and streamlining of a *Pseudomonas putida* strain**

Von der Fakultät für Lebenwissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Audrey Leprince
aus Créteil, Frankreich

1. Referent: Professor Dr. Kenneth N. Timmis
2. Referent: Professor Dr. Vitor A. P. Martins dos Santos
eingereicht am: 11.08.2010
mündliche Prüfung (Disputation) am: 05.11.2010

Druckjahr 2010

## Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

### Tagungsbeiträge

Leprince, A., Nechitaylo, T., de Lorenzo, V., Timmis, K.N. and Martins dos Santos, V.A.P.: *Pseudomonas putida* KT2440: genome minimization project. (Oral presentation, Poster). ESF-UB European conference on Synthetic Biology (ESCB): design, programming and optimization of biological systems, Sant Feliu de Guixols (2007).

Leprince, A., de Lorenzo, V., Timmis, K.N. and Martins dos Santos, V.A.P.: *Pseudomonas putida* KT2440: genome minimization project. (Poster). ESF-UB ESCB II, Sant Feliu de Guixols (2009).

Leprince, A., de Lorenzo, V., Timmis, K.N. and Martins dos Santos, V.A.P.: Streamlining the *Pseudomonas putida* genome genome for biocatalysts à la carte. (Poster). 3rd international PhD Symposium, Braunschweig (2009).

Leprince, A., de Lorenzo, V., Timmis, K.N. and Martins dos Santos, V.A.P.: Streamlining the *Pseudomonas putida* genome genome for biocatalysts à la carte. (Poster). 3rd international PhD Symposium, Braunschweig (2009). Pseudomonas 2009 XII. International Conference, Hannover (2009).

# Acknowledgements

I would like to express my gratitude to all the people who participated in one way or another to the success of my PhD work.

First of all I would like to thank my direct supervisor Prof. Dr. Vitor A. P. Martins dos Santos for giving me the opportunity to discover a new scientific field and to improve constantly my knowledge and skills on this highly interesting topic. He offered me considerable research facilities to sustain my work and the sufficient guidance to always go one step further towards the final goal.

I am grateful to Prof. Dr. Kenneth N. Timmis for the valuable discussions and guidance that he provided me since the beginning of the Thesis and for his precious help and support during the hardest moments. At the same time, I would like to thank all the members from the UMW group who always welcomed me warmly.

I take also the opportunity to extend my deepest sense of gratitude to Prof. Dr. Victor de Lorenzo. Since the first meeting he always offered me priceless advices, constructive suggestions and fruitful discussions in Spain and Germany, even when he had just a few moments free. He made me believing that everything looks simple and possible, and even though it is not always the case, he gave me the desire to always go further and never give up. During the month spent in Madrid, I had the pleasure to learn and improve new techniques but also to know wonderful colleagues, Aitor, Jose Ignacio, Rafael; Alejandro, Katharina, David, Belén, Max, Carlos, Sofía and Esther.

I would like to say “Thank You” to the colleagues who never hesitated to help me, Agata, Amélia, who started with me in the lab and participated actively to the healthy atmosphere, Andreia, Antje, Bianka, Blair, Caroline, Carolyn, Christoph, Guru, Jacek, Jörg, Joost, María, Thorben and Sylvie. Dr. Andrew Oxley and Dr. Melissa Wos-Oxley enriched my English vocabulary by taking part actively in the correction of my Thesis. I hope I won’t deceive them too much with the writing of the acknowledgment part (smiley). I would like to mention particularly among my colleagues and friends Dr. Sagrario Arias Rivas, for all the scientific discussions, helpful advices regarding different practical problems and the constructive corrections. She always gave me the impression that no task is complicated provided it’s seen in a different way. Also Dr. Mònica Bassas Galia contributed to the removal of the stress (smiley) caused by the PhD. She always helped me in my work by pointing out details missing in my presentations, posters and any kind of writing and she always “brought back the smile on my face” when it had disappeared.

A special thought for my colleagues Petra, my “Glücksbringer”, who brought me her experience and allowed me to make an important step towards the main results of my work, Dr. Piotr Bielecki, who accompanied me right since my first days in the lab by helping me with scientific exchanges until the

correction of my manuscript. I also would like to thank the new member of the group, Dr. Danielle Janus, who brought a new breeze of fresh air, experience and friendship, for all the time she offered me for discussions regarding concrete solutions as well as for the corrections. I am glad to know her. The team, more than the group, will be complete when I will have mentioned Ignacio, with whom I could always share the best as well as the toughest moments during my PhD and he contributed a considerable part of the friendly and appreciable atmosphere in the lab.

It is with emotion that I would like to say a special “Thank You” to the person who shares my life since four years now, Bharani Kumar. Without you things would not have been the same, you brought me happiness and made me feeling stronger, which is the base of my success in the lab. I hope I could express enough my feelings to you for having helped me each single minute of this long adventure.

“The last but not the least”, qui maintenant sonne mieux en anglais qu’en français, ma famille. Je voudrais la remercier chaleureusement pour m’avoir soutenue tout au long de ma thèse, et m’avoir donné le courage d’aller jusqu’au bout. A mes parents, ma sœur, Laurent, ma tante et la petite puce et à mes grands-parents je voudrais dire un grand merci pour leur amour, leur patience et leurs encouragements.

# LIST OF FIGURES

# LIST OF TABLES

# ABBREVIATIONS

| | |
|---|---|
| aa | amino acid |
| bp | base pair |
| C source | carbon source |
| cit | citrate |
| °C | degree Celsius |
| Cm | chloramphenicol |
| col. | colony |
| dist. | distilled |
| FOA | 5-fluoroorotic acid |
| g | gram |
| gDNA | genomic DNA |
| Gm | gentamycin |
| h | hour(s) |
| $H_2O$ | water |
| IR | intergenic region |
| kb | kilo base pair |
| Km | kanamycin |
| l | liter |
| MCS | multiple cloning site |
| ml | milliliter |
| $\mu_{max}$ | maximum specific growth rate |
| NA | nalidixic acid |
| ng | nanogram |
| nm | nanometer |
| OD | optical density |
| Pip | piperacillin |
| pmol | picomol |
| Rif | rifampicin |
| SB | synthetic biology |
| Tel | tellurite |
| Tc | tetracycline |
| U | unit |
| µl | microliter |
| ura | uracil |
| $x^R$ | resistant phenotype to compound x |
| $x^S$ | sensitive phenotype to compound x |

# 1. Introduction

## 1.1 The diversity of the genus *Pseudomonas*

*Pseudomonas* is a wide genus belonging to the class of gamma proteobacteria. It regroups different species exhibiting the same general morphological characteristics like rod-shaped of a length of 1.5 to 5.0 µm. These Gram-negative bacteria are motile due to the presence of one or more polar flagella. They live in aerobic milieu using a strictly respiratory type of metabolism with oxygen as final acceptor of electrons. Furthermore, they can be oxidase positive or negative but are typically catalase positive. Such properties belonged to the description of the genus present in the new edition of *Bergey's Manual of Systematic Bacteriology* (2001), characterizing *Pseudomonas aeruginosa* (Schroeter 1872) Migula 1900. This description was shown as being exhaustive by other publications which highlighted the lack of information regarding the genetic, physiologic and metabolic diversities of the genus, providing to different species the ability to live under diverse environmental conditions, utilizing a broad range of substrates (Clarke, 1982). Some *Pseudomonas* species, such as *P. putida*, were studied for their properties to grow on and degrade Halogenated Alkanoic Acids (HAA) by synthesizing dehalogenase enzymes (Slater *et al.*, 1979; Hardman and Slater, 1981). Furthermore, *Pseudomonas* species are found for example in association with the phyllosphere and rhizosphere (Lugtenberg *et al.*, 1999; Lugtenberg *et al.*, 2001), and some are threatening pathogens such as *P. aeruginosa*, which is also one of the best characterized species from the genus (Bodey *et al.*, 1983). Some other species are also well studied due to their resistance to several kinds of compound such as antibiotics and heavy metals. As already reviewed in the late 1970s, many resistances and catabolic properties are generally transposon- and plasmid-encoded, such as TOL and CAM, (Chakrabarty, 1976; Ramos *et al.*, 1997). The TOL plasmid harbors all the genetic information being necessary for the mineralization of toluene and xylene. The CAM plasmid contains the operon encoding enzymes responsible for the degradation of D-camphor. Finally, several species have been shown to be an efficient host for heterologous gene expression (Cases and de Lorenzo, 1998; Gilbert *et al.*, 2003). This quality allowed these species to be used as cell factories for the production of fine chemicals (Wubbolts and Timmis, 1990; Schmid *et al.*, 2001), of polyhydroxyalkanoates (Olivera *et al.*, 2001), for the bioremediation of polluted sites (Timmis *et al.*, 1994) and for an increasing number of other biotechnological applications.

Among the different species from the genus *Pseudomonas*, one of the most-studied and best characterized species is *P. putida*, which is able to colonize most temperate soils and waters, is non virulent and which exhibits important catabolic properties (e.g. degradation of aromatic compounds (Reineke, 1998)).

### 1.1.1 From *Pseudomonas putida* KT2440 to *Pseudomonas putida* TEC1 mutant

#### 1.1.1.1 General characteristics

*Pseudomonas putida* KT2440 (DSM6125) is a certified biological safety strain, ubiquitous and metabolically versatile soil bacterium whose genome was fully sequenced and annotated in 2002 (Nelson *et al.*, 2002). The bacterium is a plasmid-free derivative of a toluene-degrading strain, originally named *Pseudomonas arvilla* (mt-2) (Kojima *et al.*, 1967) and further reannotated as *Pseudomonas putida* mt-2 (DSM3931) (Williams and Murray, 1974; Nakazawa, 2002). It has lost the TOL plasmid and thus, the ability of degrading toluene/xylene via the *meta*-cleavage pathway (Bayley *et al.*, 1977; Franklin *et al.*, 1981). However, as other mt-2 derivative *P. putida* KT2440 retains its ability to degrade benzoate via the chromosomally encoded β–ketoadipate (ortho-cleavage) pathway, which is more detailed by Jimenez and colleagues (Jimenez *et al.*, 2002; Jimenez *et al.*, 2004). In this study they were able to highlight and analyze four main pathways for the catabolism of central aromatic intermediates corresponding to the previously mentioned β–ketoadipate pathway (with its two branches, i.e. protocatechuate and catechol), the homogentisate and the phenylacetate pathways.

The sequencing of the *Pseudomonas putida* strain KT2440 revealed a single circular chromosome of 6,181,863 bp carrying 5,420 Open Reading Frames (ORFs). The strain KT2440 is a fast growing bacterium which is able to grow on minimal medium supplemented with diverse nutrients. The global features of the genome were analyzed by Weinel and colleagues, studying the variation of the G+C, as well as the di- and tetranucleotide content. On average, the GC content which represents the relative amount of G and C nucleotides in the whole genome, reaches 61.52 % and the GC skew, which indicates within a single value the abundance of one nucleotide in comparison with the other one for a specific strand, was found to be long (0.066 on average). The whole study revealed that 80 % of the genome shares the same GC contents and oligonucleotide bias; the other 20 % are composed of 105 gene islands with atypical oligonucleotide signature (Weinel *et al.*, 2002).

After sequencing and gene annotation, the composition of the genome was revealed as follows: 184 encoded proteins which are related to mobile elements such as transposases (82 genes) and group II introns (eight genes), 51 pseudogenes (e.g. disruption by a mobile element, authentic frameshift), 804 copies of a species-specific 35 bp Repetitive Extragenic Palindromic sequence (REP) and three bacteriophage genomes. Furthermore, gene islands were found and presumably acquired by horizontal gene transfer. Some of those encode functions related with arsenate resistance, heavy metal resistance and oxidative stress response. In addition, 598 proteins were annotated as hypothetical, which means that no similarity was found with any orthologous and paralogous gene. A map of the circular chromosome was created with the different features and main role categories, see Figure 1-1.

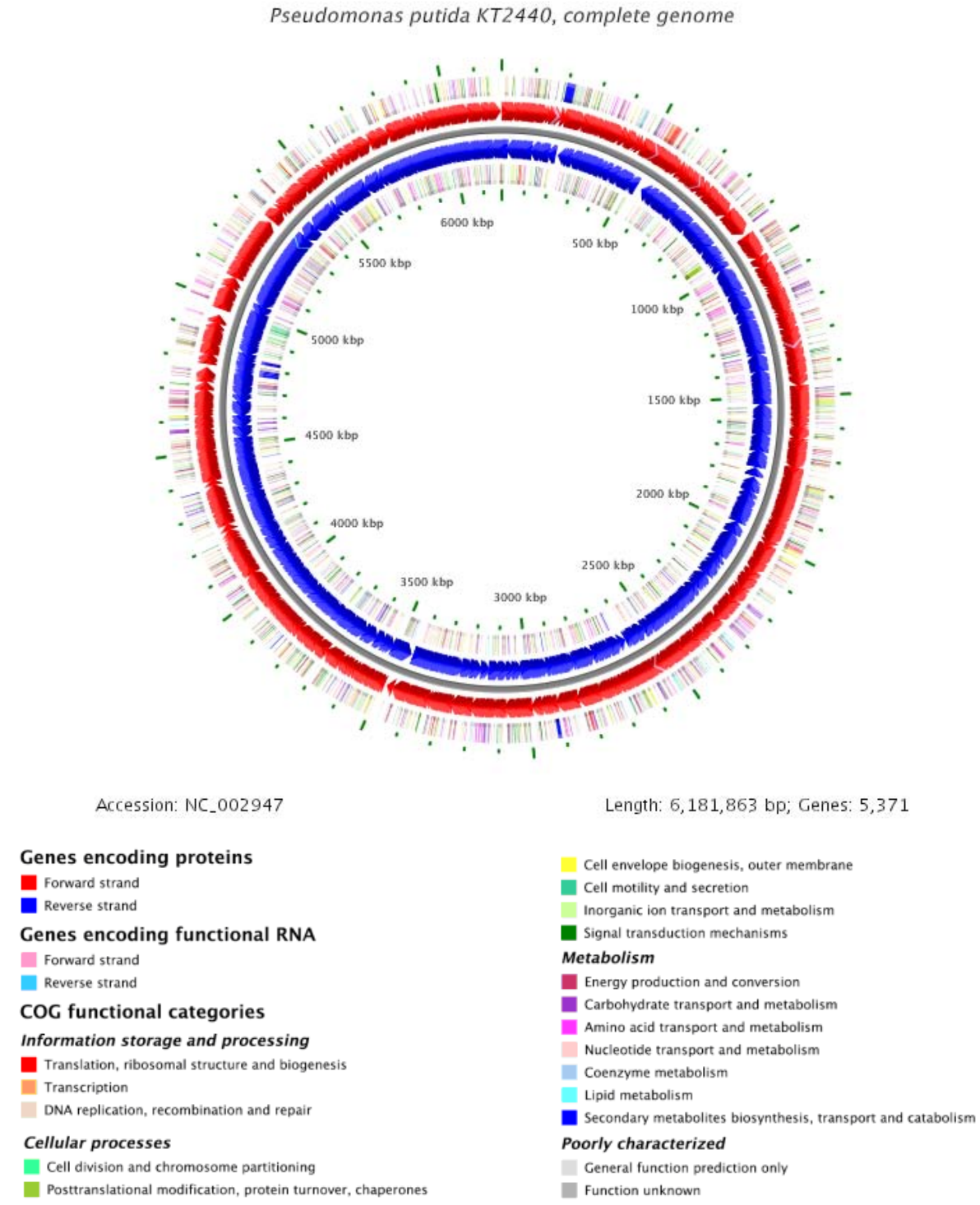



**Figure 1-1: Map of the circular chromosome of *Pseudomonas putida* strain KT2440**
(Stothard *et al.*, 2005)

In total twenty role categories were assigned to the coding sequences of the chromosome (http://cmr.jcvi.org/cgi-bin/CMR/shared/GetNumAndPercentGenesInARole.cgi; CMR), as described in Figure 1-2. The colors used for each category of the bar graph will be maintained along the thesis.

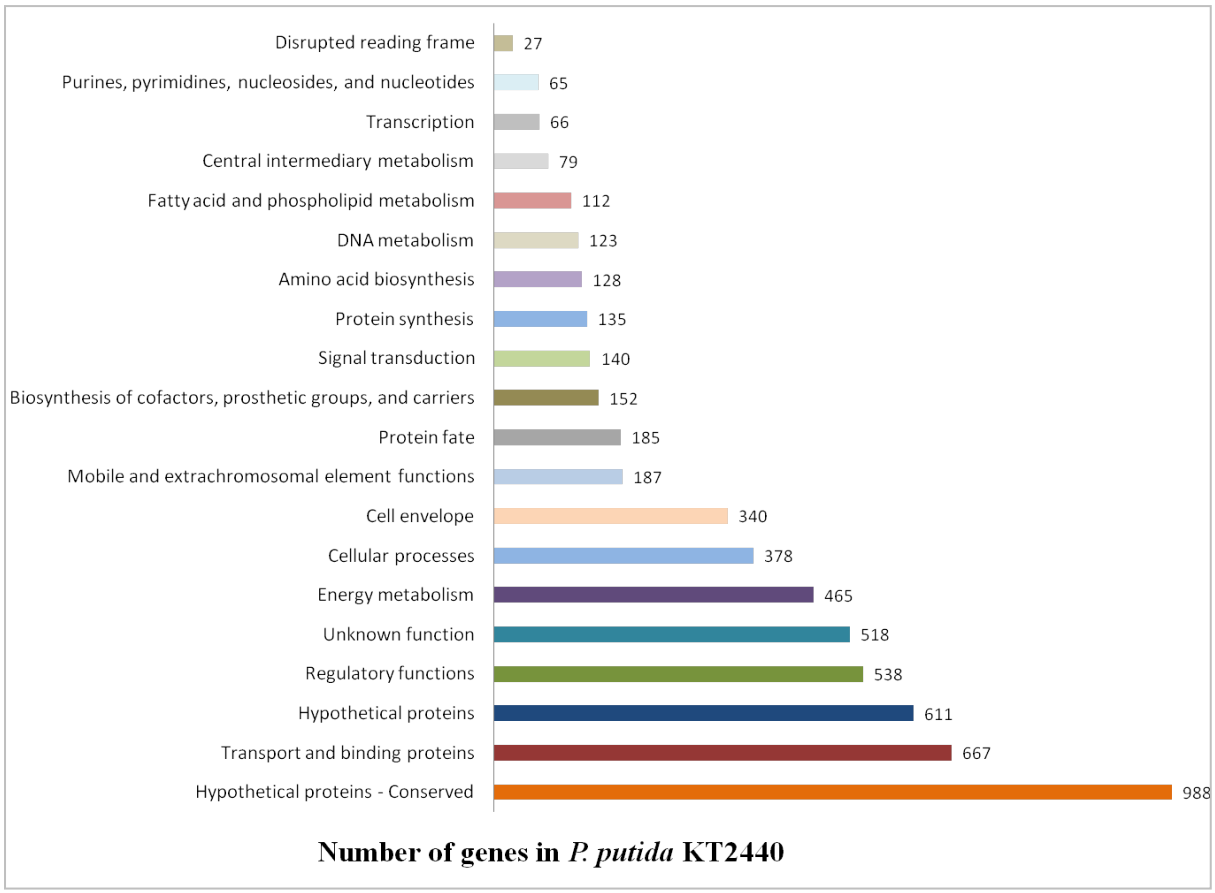



**Figure 1-2: Representation of the cellular role categories assigned to encoded proteins in *P. putida* KT2440**
The bar chart distributes the genes of *P. putida* KT2440 according to the role category assigned to the proteins they code for. One colored bar represents one category. The number of genes composing this category is annotated beside the bar. The color of the sector corresponds to the one given in the legend on the right side. In total 20 cellular role categories are found in the genome of strain KT2440 (19 cellular roles and one supplementary category for disrupted reading frames).

### 1.1.1.2 The strain as host for expression of heterologous genes

Some studies carried out in the past with different *P. putida* strains included the use of mt-2 and KT2440 as good hosts for the expression of heterologous genes (Ramos *et al.*, 1987; Cases and de Lorenzo, 1998). In 2003, Gilbert and colleagues engineered the strain by inserting a replicative expression plasmid bearing one gene for detoxification of the insecticide parathion. This gene was chosen to grow KT2440 within a dual-species consortium. They demonstrated successfully that this consortium allowed an *Escherichia coli* strain to work in a complementary manner with *P. putida*

KT2440 by first hydrolysis of parathion and further mineralization of the hydrolysis product into *p*-nitrophenol (Gilbert *et al.*, 2003).

#### 1.1.1.3 The strain as a biocatalyst

The bacterium has a natural ability to degrade organic solvents, mainly aromatic compounds like p-hydroxybenzoate, protocatechuate, aldehydes and acids through the different metabolic pathways used in the central metabolism, as mentioned in section 1.1. Peripheral metabolism is also present and functional in *P. putida* KT2440 which allows the strain to metabolize pollutants, including aromatic xenobiotics (Jimenez *et al.*, 2002). The necessary catabolic determinants were inherited from mobile elements acquired by horizontal transfer.

This degradation ability attracts the attention of many industrials for different applications in biotransformation, bioremediation and agriculture. The biotransformation of compounds such as toluene or benzoate into high value compounds is one of the examples for the use of the bacteria in industry (Wackett, 2003).

However, as many bacteria *P. putida* strains are complex and not fully understood yet. As a general phenomenon in the microbiological field, scientists started manipulating, deleting, replacing and, as of late, creating from scratch parts of organisms, an endeavor within the heart of the field of Synthetic Biology.

### 1.1.2 *Pseudomonas putida* KT2442 and TEC1

*P. putida* TEC1 originated from *P. putida* KT2442 , which is a rifampicin resistant derivative of *P. putida* KT2440 (De Lorenzo *et al.*, 1993). TEC1 was designed to serve as an example for the establishment of an efficient procedure leading to the generation of gene knockouts in Gram-negative bacteria. The strain was engineered in order to become auxotroph for uracil by disrupting the *pyrF* gene (PP_1815), a homolog of the *Saccharomyces cerevisiae* URA3 encoding the orotidine-5'-phosphate decarboxylase. The gene knockout was constructed by the replacement of the inner part of the sequence by the *xylE* gene (encoding catechol-2,3-dioxygenase), which is often used to measure different promoter activities and was amplified from a derivative of the *P. putida* TOL plasmid (Stein, 1992). The second property of this strain, due to the loss of a functional *pyrF*, resulted in the acquirement of resistance to 5-fluoroorotic acid (FOA, a uracil analogue), which in presence of the wild-type gene is transformed into a toxic compound. In parallel with this manipulation of the strain, a suicide vector (pTEC) was created carrying the wild-type version of the *pyrF* gene (Galvao and de Lorenzo, 2005). By cloning the flanking sequences of the fragment to delete and inserting in-between a specific marker into the vector, the new construct can be used for homologous recombination with the corresponding sequences of the chromosome of *P. putida* TEC1 under selective conditions (lack

of uracil, supplementation of the medium with the resistance encoded by the backbone of the suicide vector).

In summary, while conserving the general metabolic properties of *P. putida* KT2440, strain TEC1 provides the possibility of easy selection or counterselection of newly inserted genetic elements during its manipulation. It is thus an interesting candidate for the development and application of a new method aiming at the reduction of the genome.

## 1.2 Concept and tools of Synthetic Biology

The term of “Synthetic Biology” (SB) or to be exact “Biologie Synthétique” was first evoked by the book of Stéphane Leduc published in 1912, in which he described different biological processes from a physical perspective (Leduc, 1912). Sixty years later, the term “Synthetic Biology” was re-introduced by the Polish geneticist Waclaw Szybalski in 1974 (Szybalski, 1974). He wrote:

> “Let me now comment on the question "what next". Up to now we are working on the descriptive phase of molecular biology. ... But the real challenge will start when we enter the synthetic biology phase of research in our field. We will then devise new control elements and add these new modules to the existing genomes or build up wholly new genomes. This would be a field with the unlimited expansion potential and hardly any limitations to building "new better control circuits" and..... finally other "synthetic" organisms, like a "new better mouse". ... I am not concerned that we will run out exciting and novel ideas ... in the synthetic biology, in general.”

However, the concept has slowly developed only since the beginning of the 1990s. The meaning of “Synthetic Biology” evokes mainly the idea of “novelty” either based on the re-design of existing elements (like genetic circuits, enzymes…and others) or the creation and assembly of new components to create new systems and sub-systems. It was spoken of “deconstructing life” and on the other side “construction of life” (de Lorenzo *et al.*, 2006). In the past decade, the transformation of discovery-driven biology into hypothesis-driven biology made possible the control and design of new cellular functions and genetic circuits. This led to the beginning of a new era of Synthetic Biology (SB) (Benner and Sismour, 2005).

SB regroups independent yet complementary areas of research such as biology, mathematics, chemistry, molecular biology and systems biology to form a more exhaustive discipline. It is nowadays used in order to better understand and control the array of different biological systems that surround us. Two main approaches are employed by different studies and can be broadly categorized as being wet or dry laboratory based, the so-called “top-down” and “bottom-up” approaches.

### 1.2.1 The “bottom-up” approach

The so-called “bottom-up” approach consists of the fundamental reconstruction of living cells by designing and constructing new biological parts, devices and systems from non living components in conjunction with existing natural biological structures. The products of this approach are called “protocells” (Sole *et al.*, 2007). Protocell technologies have been developed at a number of research centers around the world as a way of asking questions about how it may be possible to build a synthetic organism from scratch (Rasmussen *et al.*, 2008). The basic components for a protocell assembly and simulation of life are a membrane, a promoter, a coding sequence and cell-free extract. Some complexity was brought to this concept by inserting the protein production via cascade genes into lipid vesicles. This aimed to produce protocells able to sustain biologically functional modules. For example, Toyota and colleagues designed a protocell model system based on the chemistry of oils. It takes the shape of a droplet whose metabolism (a chemical battery) is fuelled by anhydride, a substance that reacts vigorously in contact with water. The resulting system is capable of locomotion through physiochemical means as an outcome of this reaction and also exhibits a primitive form of sensation (or tropism) where the oil droplet moves towards an alkaline stimulus. This protocell model system can respond to stimuli in its surrounding environment in the form of movement that can be understood as a form of chemical computing (Toyota *et al.*, 2009).

This “bottom-up” approach is thus seen as a reproduction of the behavior of life systems. In a protocell it is aimed at combining autonomous, adaptable, robust and reusable elements. For this purpose the Registry of Standard Biological Parts was created in 2003 at the Massachusetts Institute of Technology (MIT, Cambridge, MA, USA) providing a list of formatted components (http://parts.mit.edu). Since seven years the catalog of parts and devices is constantly increasing allowing their combinations into working devices always more innovative and efficient. Each year the International Genetically Engineered Machine (iGEM) competition is organized regrouping teams from all over the world confronting on a common topic of Synthetic Biology.

In summary, the “bottom-up” tries to design new biologically functional modules whereas the “top-down” approach, which will be explained in the next paragraph, seeks to understand and re-program existing biological systems by their modification.

### 1.2.2 The “top-down” approach

The so-called “top-down” approach tends to remove existing parts, which are considered as non-essential for cellular growth and division, to reduce the biological system towards the functional essential core. This can be associated with the genomics approach seeking to understand what can be a minimal –streamlined- cell, its functions and its genome size. The term “essentiality” for a gene has to be handled with care. Being essential may involve belonging to the minimal gene set under given

environmental conditions. In contrast, the minimal gene set may not include only essential genes. So questions regarding the nature of the minimal gene set and its relationship with the essential genes under specific growth conditions remain open.

In the past decade, free living organisms and parasitic organisms, which comprise an obviously larger array of genes for survival in nature, were sequenced and compared in order to establish the first lists of potentially non-essential genes for the survival of the organism. In 1986, with initial knowledge of the yeast genomes at hand, the question of non-essential parts for cell growth and division was raised. Ten years later a hypothetical 562 kb minimal genome was inferred from *Bacillus subtilis* (Itaya, 1995). In another attempt, *Mycoplasma genitalium*, a parasitic bacterium and *Haemophilus influenza*, the first sequenced free-living organism, were analyzed in parallel to extract the so-called "minimal gene set" from the orthologous genes (Mushegian and Koonin, 1996). The interesting aspect of this comparison was due to the fact that both organisms belong to different distantly related parasitic bacterial groups in phylogeny. This analysis refers to the minimal number of genes necessary to sustain a functioning cellular life form under the most favorable conditions when the organism is able to access all the essential nutrients and is not subject to any stress condition. By comparative genomics of both strains, this minimal gene set was established and composed of a core of 256 genes (Koonin, 2000). Three years later global transposon mutagenesis was used for both *M. genitalium* and *M. pneumonia* to detect the position of genes that could be considered as essential under specific growth conditions (Hutchison III *et al.*, 1999). This approach pin-pointed 265 to 350 protein coding genes, which was very close to the amount of genes predicted by comparative genomics.

Two examples of bacteria studied for the gene essentiality are given hereafter. The analysis and deletion of genomic islands recently acquired by *E. coli* led to a >8 % reduction of its genome (under laboratory conditions), pointing to the non-essentialness of the removed genes corresponding to K-islands and transposable elements (Kolisnychenko *et al.*, 2002). The systematic inactivation of *B. subtilis* genes showed that 192 among the over 4,000 genes of the organism were indispensable and 79 other genes were predicted to be essential. About 80 % of the total amount of genes, which were determined as essential for the growth under laboratory conditions, encode for proteins acting in different categories such as information processing, cell envelope, shape, division and energetics (Kobayashi *et al.*, 2003).

These lines of progress form a basis for future development on the streamlining of genomes. Although the knowledge of individual genes and their function (essential or not) remain important, a determining factor seems to be their role in the context of the cellular networks in which they participate. By combining *in silico* comparative genomics and *in vivo* global knockout mutagenesis, most of the genes of the minimal gene set described for *M. genitalium* have been proved to be either universal or conserved in all bacteria (Forster and Church, 2006).

Until now, no *Pseudomonas* species was studied for the establishment of a potential minimal set of genes. However, recent metabolic reconstruction and genome-scale, constraint-based models of the minimized genome have been done on the basis of the genome sequence data sets, biochemical information and strain-specific knowledge (Papin *et al.*, 2003). Two metabolic reconstructions of *P. putida* KT2440 have been published proposing a set of essential metabolic functions and putative essential genes (Nogales *et al.*, 2008; Puchalka *et al.*, 2008). These new *in silico* information and the recent publication of Molina-Henares and colleagues show an increasing and more intensive work done on KT2440 to better understand the genetic, physiologic and metabolic properties of the strain (Molina-Henares *et al.*, 2010).

## 1.3 Mutagenesis

Mutagenesis is a process by which the genetic information of an organism is changed in a stable manner, either in nature or experimentally. In order to study the gene function in a given organism and gene essentiality under defined growth conditions, mutagenesis is an appropriated tool and one of the most popular. Perturbation of the whole system by simple mutation in the genome can be detected by comparison between the generated mutant strain and its corresponding wild-type under specific conditions. The mutation can be provoked by different types of procedures, some of them are mentioned below.

- The insertional mutagenesis allows additional bases to be inserted into the genomic DNA generating possible different phenotypes.
- The site-directed mutagenesis is used to disrupt a targeted gene by using in general vectors carrying the mutated gene or carrying markers inserting in the targeted gene. This procedure requires precise information about the sequence of the gene as well as a high frequency of homologous recombination.
- The transposon mutagenesis consists of inserting randomly mobile DNA elements into a host strain. This procedure on the contrary does not need any previous knowledge about the sequence of the studied strain. This case will be further described in the next paragraphs.

The study of bacterial genomes has always necessitated the use of mutagenesis tools in order to assess the gene function in-depth. Generation of mutants in bacteria is a relatively straightforward process due to easily detectable phenotypes. Several techniques were established since the beginning of the molecular biotechnology field in order to mutate target genes and therefore, to study their role and importance for a single cell. Mutation frequency can be increased by induced mutagenesis where bacterial strains are exposed to chemical or physical mutagenic agents. Such agents are represented for instance by different chemical compounds or X-rays. In general, these procedures require an important screening process in order to find the desired mutants. A third option is based on genetic

techniques which proved in the past their efficiency for the determination of gene functions. The allelic exchange generating the disruption of a target gene, the use of antisense RNA molecules for inhibition of the transcription and finally the transposon mutagenesis were the three main techniques developed to generate mutants. For all of them a common tool was developed consisting of a series of vectors designed and assembled to bring the mutation or to force the mutation to occur within the chromosome of a bacterial strain. Vectors are circular DNA molecules specifically designed for a precise aim such as site-directed mutagenesis or transposon mutagenesis in a given genus or species. The cloning vectors are on average small (between 2.5 and 5 kb) to assure a better propagation in the host strain. The main features which have to be considered during the design and which give the specificity to the vector concern the following elements:

- An origin of transfer is necessary for transferring the vector from the donor strain to the host by conjugation. The origin is encoded by the *tra* genes, but the site where the transfer of the vector starts is called *oriT*. For self-transmissible vectors all the genes necessary to the transfer are present in the same strain; however, often mobilizable vectors, used for mutagenesis, do not contain these specific genes. This situation leads to the use of transfer by triparental mating involving the donor strain with the mobilizable vector, a helper strain carrying the *tra* genes, and the host strain.
- An origin of replication is preferably designed as a conditional origin which will provide the replicative ability to the vector in the donor strain but not in the host strain. After its transfer to the host, the vector behaves as a suicide plasmid which means that is forced to integrate the host chromosome.
- A Multiple Cloning Site (MCS) will offer to the user a panel of restriction sites, available for the cloning of the fragment of interest (e.g. mutated gene or resistance cassettes).

### 1.3.1 Transposon mutagenesis

Transposon mutagenesis which provides wide collections of knockout mutants plays a key role in the determination of gene function and design of regulatory networks. It is based on the random insertion of transposable elements, in general genetically engineered. These mobile DNA sequences were classified in different ways, but always separating transposons and insertion sequences (Craig, 1996; Mahillon and Chandler, 1998). They are recognizable from each other due to their variation from one element to the other in regard to their size, structure, transposition mechanism and target site. Bacterial transposons were discovered in the 1970s. Two main categories regroup the transposable elements in bacteria: insertion sequences (IS) and transposons (Tn). Insertion sequences, simpler in length and organization (between 700 and 1,300 bp), are composed of two inverted repeats flanking the transposase gene being responsible for delocalization of the mobile DNA within a molecule into a different site. On the contrary, transposons are more complex in their composition. Two IS, in direct

or inverse orientation, flank genes coding for antibiotic and heavy-metal resistance or catabolic functions. A specificity of the transposons involves their random insertion at multiple sites of a chromosome. Another characteristic of the transcription is that transposable elements insertion can affect the transcription of the genes located in the vicinity of the transposon. This depends on the position of the insertion in an operon, e.g. if the disrupted gene is at the beginning of an operon, and the composition of the transposon, e.g. with an internal transcription terminator (Gerdes *et al.*, 2002). Transposable elements were discovered due to this property more than 60 years ago. Due to all these characteristics, transposons are the most often used tool for wide-genome mutant library generations.

Transposon mutagenesis is a powerful tool when it is used either to transfer different parts of the chromosome from one cell to another or to insert foreign DNA into a target gene. *In vivo* and *in vitro* mutageneses were developed with both their advantages and disadvantages. Briefly, by using *in vitro* transposition, which involves the incubation of the target DNA with transposable molecules and purified transposase, high saturation levels are reached. Furthermore, different loci of DNA targets can be focused on depending on the aim of the mutagenesis. However, this method is labor-intensive and requires initial information related to the target DNA.

The *in vivo* mechanism involves the transfer of the transposon-harboring suicide delivery vector to the host chromosome by transformation (electroporation) or conjugation. The best conjugative transposable elements studied so-far are Tn*916* (the smallest one, 18.4 kb), and Tn*1545*, extracted from *Enterococcus* and *Streptococcus* species (Gawron-Burke and Clewell, 1982; Courvalin and Carlier, 1986). This system of delivery does not require a naturally competence of the host to insert these transposable elements. Disadvantages can be found in the preliminary steps concerning the generation of the delivery system namely the cloning of the transposable elements in a suicide vector and the choice of a transposase which should be expressed in the host. Finally genetic tools have to be employed in order to remedy the insertion instability. This latter characteristic is due to the transposase which usually provokes a secondary transposition (Goryshin and Reznikoff, 1998). Different solutions were found in the past to provide improved system for *in vivo* mutagenesis. All the experiments of the PhD work is also carried out *in vivo*.

One of the most studied transposons is the Tn*5*, a composite transposon within the IS*4* family of mobile elements, which was used for both *in vitro* and *in vivo* mutagenesis. Its particular structure binds the transposase to the specific ends of the transposon, called synaptic or paired end complex. The simplicity of its organization and randomness makes it one of the most utilized transposons for genetic rearrangement (Reznikoff, 2003). The structure of the Tn*5* is extended to a 5.8 kb sequence composed of three antibiotic resistance markers framed by two inverted 1.5 kb IS*50* insertion sequences. Kanamycin, bleomycin and streptomycin resistances are expressed by a promoter situated in one of the adjacent IS sequences (IS*50*L). Genetic information specific to the Tn*5* transposon is

contained in the IS*50* elements, IS*50*L and IS*50*R. Both of them encode proteins used for the transposition process of the whole element, the transposase (Tnp) and the transposition inhibitor (Inh). Each of the IS*50* elements is framed by two 19 bp *cis*-active sites, the outside Oend (OE) and the inside Iend (IE) parts, see Figure 1-3. These short sequences are recognized by certain proteins of the host organism which thus can regulate the transposition action. The two sequences are differing from each other at a single position providing them different roles during the transposition.

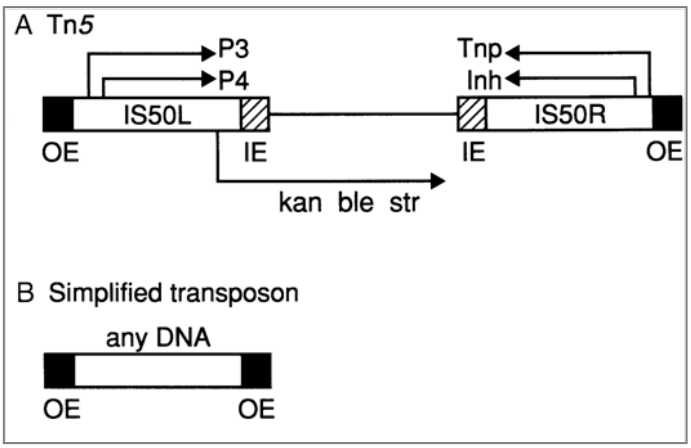



**Figure 1-3: Organization of the Tn5 transposon**
**A.** The ends are indicated as OE and IE framing both IS*50*L and IS*50*R insertion sequences. IS*50*R encodes for both the transposase (Tnp) and the inhibitor (Inh). **B.** A simplified transposon is represented with the sufficient and necessary features to allow the binding to the DNA for the formation of the synaptic complex. (Source: modified from Reznikoff, 2003)

As mentioned before, the transposon is a complex structure. However this complexity provides all the necessary tools for an autoregulation of its transfer to a target DNA. As an example the Tnp and Inh proteins are chosen to described the mechanism of regulation. From the one hand the transposase acts mainly in *cis* (Johnson *et al.*, 1982) which implies that only transposons encoding the transposase or situated close to the DNA encoding it, will be mobilized by the Tnp (Figure 1-4). After mobilization, the Tnp binds to the ends and dimerizes. The Tnp is responsible for the contact between two dimerized transposases leading to synaptic complex formation and for the dimerization of the Tnp with the target DNA to release this complex). The action of the Tnp in *cis* should provoke a low frequency of the recombination. On the other hand the second protein (Inh) inhibits the transposase by their dimerization with Tnp which leads to the inability to form the synapse. These two elements of regulation are just two examples among a larger group not mentioned here but well described in literature (Reznikoff, 2003).

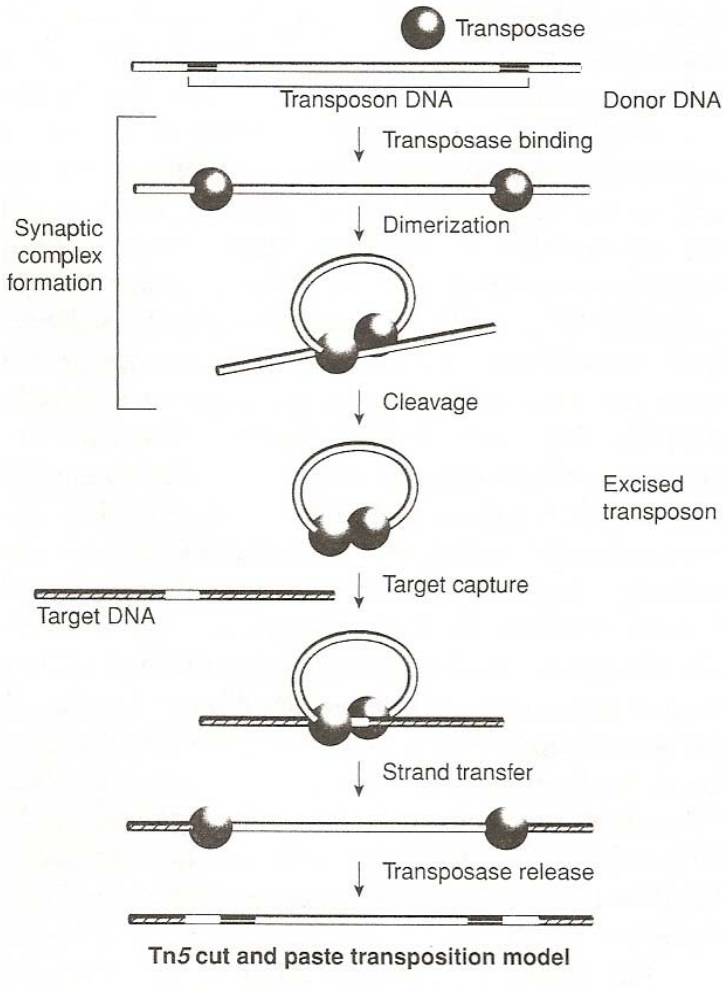



**Figure 1-4: Tn*5* transposition from a donor DNA to a host organism by cut-and-paste conservative mechanism**
The transposase binds to the end sequences of each IS*50* element leading to the formation of the synapse by oligomerization of the end-bound transposase monomers. The blunt end cleavage of the transposition synaptic complex from the donor molecule precedes its binding to the target DNA and transfer of the transposon 3' end to a 9-bp target sequence. The Tnp is finally released from the transposition product *in vivo* and the 9-bp gaps are repaired at either end (Reznikoff, 2002).

### 1.3.2 Mini-transposons

As mentioned above, transposons are composed of different recognition sequences which make them complex. This handicaps the user for their handling. Therefore, mini-transposons as well as their delivery systems were designed and developed in order to keep only the essential parts for transposition and recognition and to guarantee an efficient transfer to the chromosome. In order to provide a wide choice for the potential user the Tn*10* and Tn*5* mini-transposon derivative series were created with different markers (De Lorenzo and Timmis, 1994). They were engineered from the Tn*5* and Tn*10* transposon families. Consequently, they conserved their ability to transpose in a conservative way. This means that the whole transposon or mini-transposon will be transferred to the target site but lost in the donor molecule.

The coding sequence for different antibiotics like kanamycin, tetracycline, streptomycin-spectinomycin and chloramphenicol was incorporated in the design of the mini-transposons (De Lorenzo *et al.*, 1990). However, for some purposes the acquisition of antibiotic resistance by the engineered strains was not appropriate due to the utilization of the strains for further experiments or industrial purposes. Therefore other markers have been used like resistances to chemical compounds such as mercuric salts or potassium tellurite and other reporter genes (e.g. *lacZ*, *phoA*) (De Lorenzo *et al.*, 1990; Herrero *et al.*, 1990; Sanchez-Romero *et al.*, 1998). The mini-Tn*10* and mini-Tn*5* were borne by the high-copy conditionally replicative R6K plasmids (Kolter, 1981) in order to facilitate their insertion in bacterial chromosomes due to their suicidal behavior in the absence of the π protein. Moreover, a mutant version of the transposase coding gene from IS*50*$_R$ was obtained by removing the single *Not*I site from its sequence and placed directly outside the mini-Tn*5* end. Insertion of the mini-transposon derivative itself in the genome could therefore occur and the rest of the plasmid containing the transposase gene and its inhibitor was lost. One important point was that further rearrangements of the mini-transposon through the genome were disabled but the cell remained accessible to other mobile DNA elements inserted by further transposition events. In this study the mini-Tn*5* derivative series was used to create adapted mini-transposons for random insertion and chromosomal deletion in *Pseudomonas putida* TEC1.

### 1.3.3 Mini-transposon libraries

Different mini-transposon-based mutants were created in the last decade for several bacterial strains. The study of the different roles of proteins encoded by the knockout genes under specific growth conditions as well as the discovery of new functions for these proteins represented some of the goals of the establishment of knockout mutant libraries. Some examples of mini-transposon libraries are introduced hereafter to give an overview of the work done in the past 10 years.

- The yeast *Saccharomyces cerevisiae* was studied by transposon tagging and more than 11,000 mutants were collected. The work predominantly focused on the detection of new genes and the creation of insertion alleles (Ross-Macdonald *et al.*, 1999).
- The two *Escherichia coli* strains MG1655 and DH10B were subjected to genetic footprinting for the discovery of new antimicrobial drug targets. For this purpose two mini-Tn*5* derived transposon libraries were constructed and published in 2002 (Gerdes *et al.*, 2002).
- In the same year, 1,056 knockout mutants of the *Pseudomonas aeruginosa* PAO1 strain were generated using 11 different tagged pUT miniTn*5* Km2 mini-transposons. The library was screened to identify conditionally essential genes *in vivo* (Lehoux *et al.*, 2002).
- In 2003 the first list of Tn*5* derivative transposon mutants for *P. aeruginosa* was published (Jacobs *et al.*, 2003). More than 30,000 individual insertions were obtained and mapped.

Three years later another library was established for *P. aeruginosa* PA14 with more than 34,000 mutants (Liberati *et al.*, 2006).

These mini-transposon libraries were useful for the confirmation or new annotation of specific genes, the determination of the essentiality of particular genes and the dispensability of others. Finally, the generation of a considerable amount of mutants allowed the re-construction of gene regulatory networks.

# 2. Project rationale

## 2.1 Aim of the project

The goals of this project are summarized hereafter:

- Development and improvement of a set of tools aiming at the simultaneous excision of random target genes in a *P. putida* KT2440 derivative genome, *P. putida* TEC1 strain, under specific laboratory conditions.
- Application of these tools for the generation of single and double mutants carrying either a single gene disruption or two independent disruptions in the same strain. This will lead to the establishment of a list of genes which will be confirmed to be non essential for sustaining a functioning cellular life under the given conditions and thus to a possible improvement of the knowledge about the composition of a minimal gene set.
- Reduction and simplification of the *P. putida* TEC1 genome for a better understanding and control of the metabolic and regulatory hierarchies underlying cellular processes under given biocatalysis-relevant conditions. This application prepares the streamlined genome for a re-programming of the cell (endowed with a series of highly coordinated, newly assembled genetic circuits) for further manipulations aiming at the creation of a better performing biocatalyst. Furthermore, by achieving such constructs as a proof-of-principle, it was aimed at laying the foundation for the construction and engineering of cells performing effectively and efficiently specific functions of biotechnological, environmental or medical interest.

*P. putida* was chosen due to its natural properties to colonize a wide range of milieus and mainly due to its ability to degrade aromatic compounds. Although *P. putida* KT2440 is one of the best-studied and characterized *Pseudomonas* strains, over 30 % of its genes code for proteins whose functions are not yet fully understood. This uncertainty makes it more difficult to design and re-program cellular functions. Furthermore, known proteins such as transposable elements can also provoke instability. Removal of non-essential genes reduces the consumption of energy by the cell, limits unexpected reactions that may affect the control of pathway relevant for biocatalytic applications and enables to simplify the mesh of regulatory elements that underlie cellular functions. Streamlining the genome of the bacteria is pivotal for the project illustrated in Figure 2-2, which will be described after presentation of the different tools utilized for the genome reduction in the next paragraphs.

## 2.2 Development, combinations and applications of genetic tools

The development of the tools used for the general concept of random genomic deletion was based on two main axes: (i) random mutagenesis with random insertion of genetically modified mini-transposons in association with (ii) a site-specific recombination system. (i) Transposon mutagenesis was chosen as part of the method due to its random property. Random insertion of the mini-transposon in the genome does not require any precise information about the genes to be disrupted, and further removed, which suits for working with a strain in which one third of the chromosome encodes proteins whose function is not yet determined. Mini-Tn*5* mini-transposons, introduced in paragraph 1.3.2, were engineered for a better application of the random mutagenesis to the *P. putida* TEC1 strain. (ii) The Flp-*FRT* site-specific recombination system from *Saccharomyces cerevisiae* was used as the final step for deletion of a genomic fragment from the chromosome.

For the different mini-transposons created in the past, specific single restriction sites were inserted for the cloning of new marker genes and/or supplementary specific genes of interest (De Lorenzo and Timmis, 1994). The construction of the two mini-Tn*5* derivatives used for this study was based on the different elements described in the next paragraphs.

### 2.2.1 Suicide vectors: mini-transposon delivery system

The pBAM1 and pJMT6 vectors, see Table 3-1, were used as template for the generation of the two mini-transposon derivatives. They belonged to the category of mobilizable vector, described for their use in mutagenesis experiments (Sanchez-Romero *et al.*, 1998). In this case, the fragment inserted in the MCS corresponds to the mini-transposons derivatives. They carry respectively the kanamycin resistance and tellurite resistance encoding genes.

### 2.2.2 Selectable markers for distinguishing mini-transposons

Two different resistance cassettes, kanamycin (Km) and tellurite (Tel), were chosen in order to facilitate the recognition of the insertion of each mini-Tn*5* derivative in the genome. The Km cassette was the first mini-Tn*5* manipulated. Due to efficiency problems in *P. putida* strains, tellurite was preferred to tetracycline (Tc) for inclusion in this study. The second mini-transposon was thus based on the non-antibiotic resistance cassette composed of the *kilA-telAB* genes encoding the resistance to tellurite. This mini-Tn*5* derivative was proven in the past to provide high levels of resistance to a *P. putida* KT2440 derived strain (Sanchez-Romero *et al.*, 1998).

### 2.2.3 Recombination system leading to genomic deletion

The Flp-*FRT* site-specific recombination system from the yeast *Saccharomyces cerevisiae* was first described in 1995 (Sadowski, 1995). Concrete applications for diverse bacteria were suggested eight

years later by Schweizer and colleagues who presented the system as a powerful tool for high-throughput genetic analysis (Schweizer, 2003). The system uses two directional binding sites, the so-called Flp Recognition Target (*FRT*), in combination with the flippase (Flp) protein. The main goal is to excise marker genes of experimental interest from the genome without affecting the desired phenotype. The 2 µm plasmid from *S. cerevisiae* produces naturally the recombinase, as many *Saccharomyces* yeast do. The general procedure and consequences of the action of the Flp recombinase is shown in Figure 2-1, where two different cases are shown. The recombinase always acts when two *FRT* sites are present on the same DNA molecule either in the same direction (deletion case) or in the opposite one (inversion case). The method was already applied to many organisms and recently scientists developed new engineered Flp-*FRT* system to adapt to any given filamentous fungus (Kopke *et al.*, 2010). The modification of the system aimed at generating knockout strains and recyclable antibiotic resistance markers for filamentous fungi.

In the present study the goal was slightly modified. The two recognition sites were inserted separately, involving the presence of an undefined DNA fragment in the middle. Therefore, after action of the recombinase a whole genomic fragment was deleted along with the marker genes. This means that through the random integration of the *FRT* sites in the genome (via cloning in the mini-Tn*5* derivatives), the Flp-*FRT* site-specific recombination system participated actively to a high-throughput deletion method.

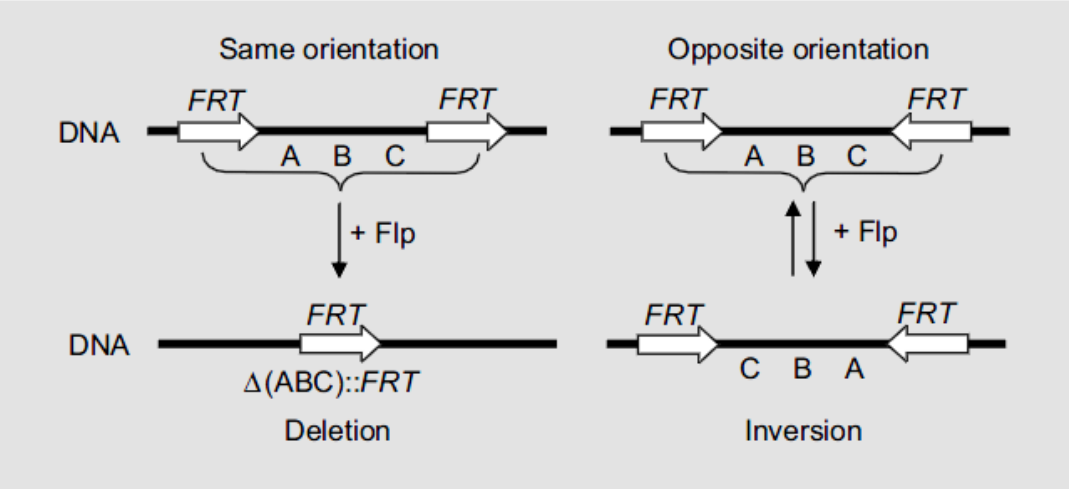



**Figure 2-1: Principle of the Flp recombinase action on two *FRT* sites in the same DNA fragment**
DNA fragments are represented on the top, carrying two FRT sites each. On the left, both *FRT* sites are inserted in the same direction, which after action of the flippase, leads to deletion of the fragment situated between both recognition targets. The deletion is symbolized by Δ and the name of the deleted genes in brackets. On the right, the *FRT* sites are inserted in opposite direction which leads to the inversion of the DNA fragment after action of the flippase. (Schweizer, 2003)

### 2.2.4 Use of the wild-type *pyrF* operon as counterselectable marker

The wild-type *pyrF* operon was cloned into one of the mini-Tn*5* derivatives in order to be inserted randomly into the genome of the *P. putida* TEC1, a uracil auxotroph mutant of *P. putida* KT2440 (paragraph 1.1.2). It was mainly used as a counterselectable marker during the process of deletion due to the sensitivity of the single and double *P. putida* TEC1 mutants (with one or two mini-Tn*5* derivatives inserted in the genome) to 5-fluoroorotic acid. By inserting the mini-transposons, the wild-type operon was recovered by the chromosome. As it will be described in more details in section 4.2.1, the genomic deletion was planned to remove the different markers carried by both mini-Tn*5* derivatives and leave behind a single *FRT* site framed by one end of each mini-transposon. Therefore the simultaneous presence of the operon and the chemical compound would in theory force the deletion to occur for the cells to survive.

The general scheme of the experiments carried out during the PhD work leading to the deletion of random fragments in a single *P. putida* TEC1 strain is presented in Figure 2-3. The phenotype of each wild-type and intermediary mutants, which leads to the deletion of a genomic fragment, is annotated on the scheme (green). Each step of the procedure (in blue) is shortly described indicating the type or reaction and the nature of the inserted element. The steps 1. and 2. utilize the two different mini-Tn*5* derivatives carrying a single *FRT* site each. At the bottom of the scheme, the deleted mutants are found and the *P. putida* $\Delta_x$ precises the number of deletions in the mutant. This whole procedure will lead to the reduced cell shown at the right side of Figure 2-2, in which the several rounds of deletion can be visualized. In the blue-stacked Venn diagram each new round of genomic fragment removal is represented by a smaller circle. The red rectangles that act for the *FRT* sites show the increasing number of possibilities for recombination with the decrease of the genome size.

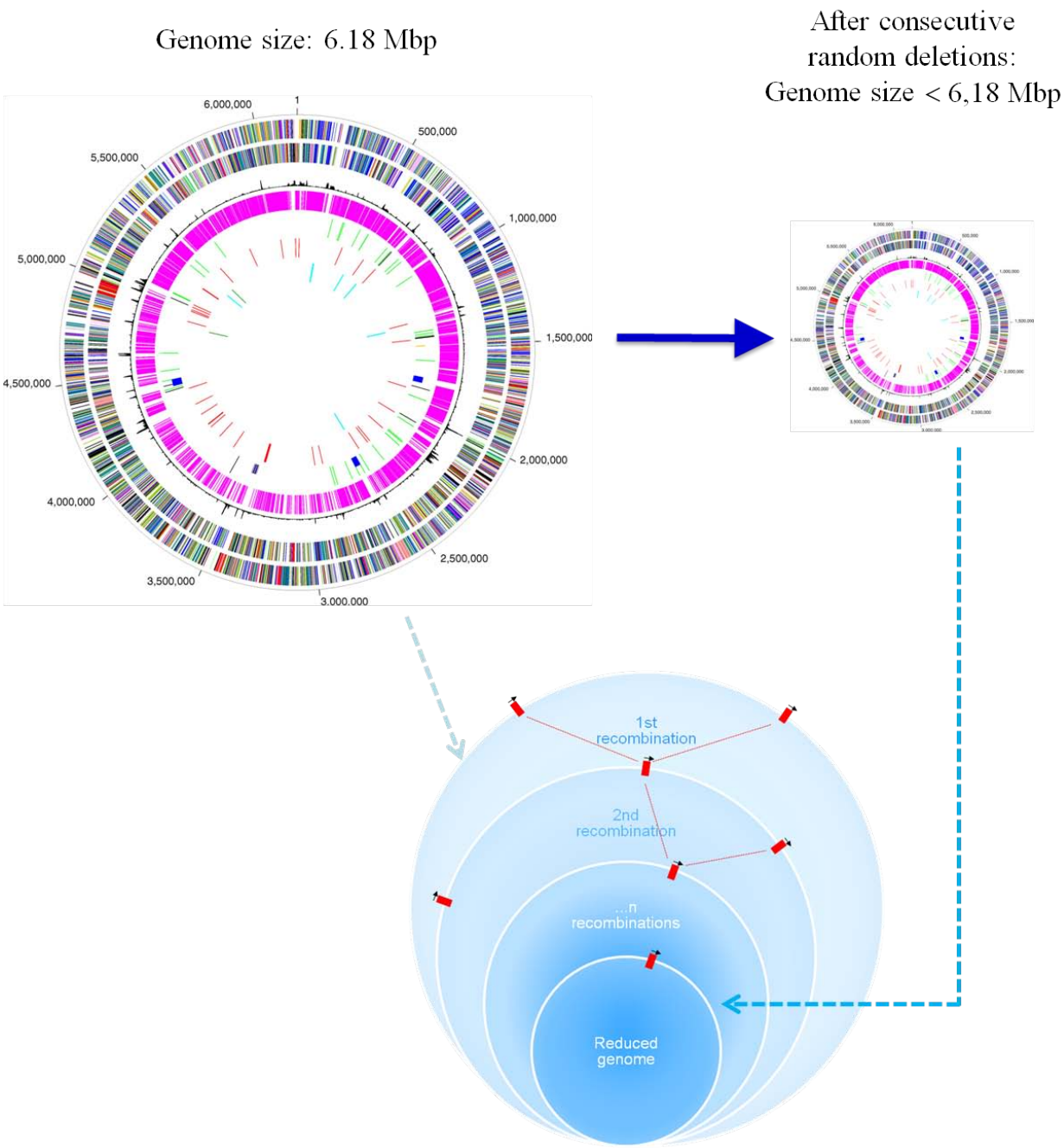



**Figure 2-2: Illustration of the genome reduction concept applied to *P. putida***
After iteratively applying the developed procedure for random genomic deletion in *P. putida* TEC1, the genome (originally around 6.18 Mbp) is reduced. However, due to the random aspect of the method, the length of the deleted fragments cannot be predicted in advance. The genome reduction is represented by the minimization in size (dark blue arrow) of the picture from the circular chromosome of strain KT2440 (Source: K.E. Nelson *et al*). At the bottom of the Figure the physical reduction of the length of the chromosome is schematized by a stacked Venn diagram in which each smaller circle represents a new deleted mutant. Each reduced genome becomes a new host for insertion of mini-transposon derivatives which leads to a new round of deletion and a further reduced genome. The light blue arrows indicate at which step the circular representations of the chromosome (top of the Figure) correspond. The red rectangles stand for one FRT site and are linked by red lines indicating the recombination and deletion event. The number of reduction rounds is indicated in the different discs of the diagram.

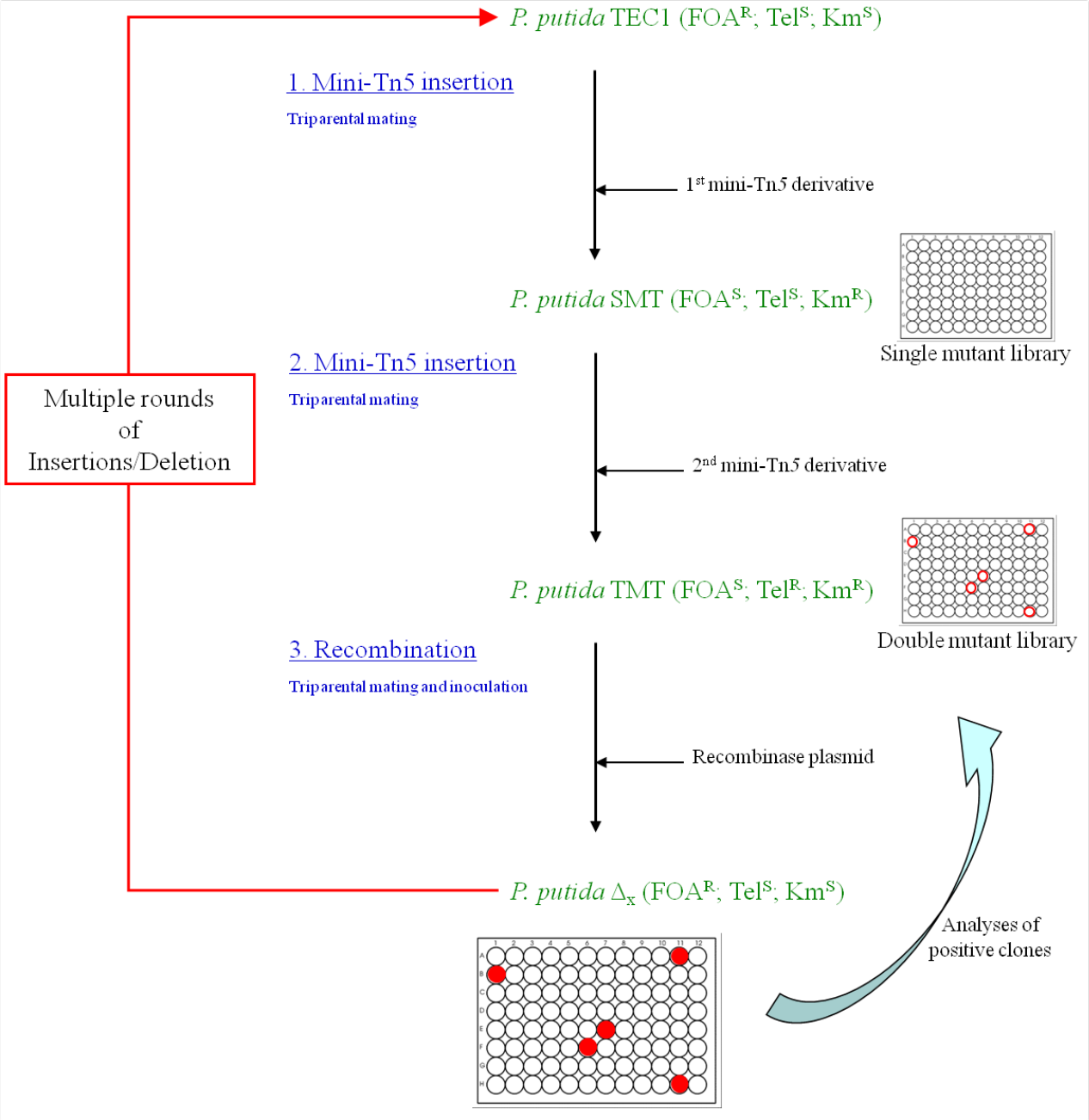



**Figure 2-3: General representation of the procedure for consecutive random genomic deletion in a single strain**

The different steps, from the first insertion of minitransposon derivative into the genome of the *P. putida* TEC1, considered as the wild-type strain, until the first random genomic deletion, are highlighted following a vertical scheme. The names of the different mutant strains generated with this procedure (green) describe the amount of mini-Tn*5* derivative inserted in the genome. **SMT** stands for Single MiniTransposon. **TMT** stands for Two MiniTransposon. $\mathbf{\Delta_x}$ indicates a genomic deletion (**Δ**) and the number of rounds of deletion (**x**). The generation of the SMT and TMT strains leads to the construction of single and double mutant libraries, respectively. After each final step of genomic deletion the positives clones are analyzed to obtain more insights into the characteristics of the deletion.

# 3. Materials and methods

## 3.1 Bacterial strains and plasmids

The different bacterial strains and plasmids employed during the PhD work are listed in Tables 3-1 and 3-2 below.

**Table 3-1: Bacterial strains used for the generation of randomly deleted mutants**

| Strains | Relevant characteristics | Reference |
|---|---|---|
| *E. coli* CC118λ*pir* | Δ(*ara-leu*) *araD* Δ*lacX74 galE galK phoA20 thi-1 rpsE rpoB argE recA1* λ*pir* lysogen | (Herrero *et al.*, 1990) |
| *E. coli* HB101 | $Sm^R$; *rpsL recA thi pro leu hsdR*$^-$ *M*$^+$ | (Sambrook *et al.*, 1989) |
| *E. coli* DH5α | F$^-$ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^-$ $m_K^+$), λ– | (Meselson and Yuan, 1968) |
| *P. putida* KT2442 | $Rif^R$; derived from *P. putida* KT2440 | (De Lorenzo *et al.*, 1993) |
| *P. putida* TEC1 | $Rif^R$; *P. putida* KT2442 with a nearly complete deletion of *pyrF* | (Galvao and de Lorenzo, 2005) |
| *P. putida* SMT | $Km^R$; *P. putida* TEC1 with a mini-Tn*5 KpF* randomly inserted in the chromosome | This work |
| *P. putida* TMT | $Km^R$; $Tel^R$; *P. putida* SMT with a mini-Tn*5 TF* randomly inserted in the chromosome | This work |

**Table 3-2: Plasmids used for the creation of randomly deleted mutants**

| Plasmids | Relevant characteristics | Reference |
|---|---|---|
| patt*FRT* | $Ap^R$; pattTn7 with a *FRT* sequence from pPS856 | (De las Heras *et al.*, 2008) |
| patt*FRT-pyrF* | $Ap^R$; delivery vector for < *FRT-pyrF*> cassette | This work |
| pBAM1 | $Ap^R$ $Km^R$; delivery vector for mini-Tn*5 Km* | Martínez García *et al*, submitted |
| pBAM1/*KpF* | $Ap^R$ $Km^R$; delivery vector for mini-Tn*5 KpF* | This work |
| pBBFLP | $Tc^R$; delivery vector for FLP recombinase | (De las Heras *et al.*, 2008) |
| pBBR1MCS | $Cm^R$; broad-host-range vector, *ori*ColE1 *mob*RK2 | (Kovach *et al.*, 1995) |
| pBBR1MCS-5 | $Gm^R$; replacement of $Cm^R$ cassette by $Gm^R$ cassette from the pBBR1MCS vector | (Kovach *et al.*, 1995) |
| pJMT6 | $Ap^R$ $Tel^R$; delivery vector for mini-Tn*5 Tel* | (Sanchez-Romero *et al.*, 1998) |

| | | |
|---|---|---|
| pUT | $Ap^R$; *tnp** gene of Tn*5*-$IS50_R$ inserted in *Sal*I site of pGP704 | (Herrero *et al.*, 1990) |
| pUT/*TF* | $Ap^R$ $Tel^R$; delivery vector for mini-Tn*5* *TF* | This work |
| pRK600 | $Cm^R$; *ori*ColE1 *mob*RK2 *tra*RK2 | (De Lorenzo and Timmis, 1994) |

## 3.2 Medium and supplement

### 3.2.1 LB (Luria Bertani) medium

- ***Escherichia coli* strains**

LB medium was prepared as a general stock for the different *Escherichia coli* strains and supplemented when necessary with the appropriated antibiotics and other chemical compounds. The basic recipe was taken from the Molecular Cloning manual from (Sambrook *et al.*, 1989) and described hereafter:

| | |
|---|---|
| Bacto-trypton | 10 g |
| Yeast extract | 5 g |
| NaCl | 10 g |
| dist. $H_2O$ | filled up to 1 l |

The pH of the LB Broth was adjusted to 7.5 with NaOH and filled to 1 l. For the preparation of LB Agar plates 15 g of Bacto Agar were added to 1 L of LB Broth before sterilization by autoclaving.

- ***Pseudomonas putida* strains**

The LB medium prepared for the different *P. putida* strains generated in this work was based on the same recipe previously described but supplemented with citrate (cit, 0.2 %) and uracil (ura, 20 µg/ml). The uracil was added in the case of the wild-type strain and the deleted mutant strains lacking the *pyrF* operon as explained in section 1.1.2. The *P. putida* SMT (Single Mini Transposon) and TMT (Two Mini Transposon) strains did not require its presence in the medium.

### 3.2.2 Φ Broth medium

The Φ Broth medium was only employed for the preparation of *E. coli* competent cells. The following recipe was used:

| | |
|---|---|
| Bacto-trypton | 10 g |
| Yeast extract | 2.5 g |
| KCl | 3.75 g |

| | |
|---|---|
| $MgSO_4.7H_2O$ | 4.05 g |
| dist. $H_2O$ | filled to 1 l |

Before adding the water until 1 l, the pH was adjusted to 7.6. Sterilization was done by autoclaving.

### 3.2.3 Minimal medium

In order to prepare the M9 minimal medium, a stock of M9 salts (10x concentrated) was first reconstituted following the instructions of the Molecular Cloning manual (Sambrook *et al.*, 1989). To 800 ml of distilled water the following chemical components were added:

| | |
|---|---|
| $Na_2HPO_4.2H_2O$ | 85 g |
| $KH_2PO_4$ | 30 g |
| NaCl | 5 g |
| $NH_4Cl$ | 10 g |

The pH was adjusted to 7.0-7.2, the volume was filled to 1 l, and the solution was autoclaved.

To prepare 1 liter of M9 minimal medium the following sterile components were mixed:

- 100 ml of M9 salts (10x concentrated)
- 1 ml of 1 M $MgSO_4$ (filter-sterilized)
- 10 ml of 20 % citrate (filter-sterilized)
- 889 ml sterile dist. $H_2O$

To prepare solid M9 agar plates, 7.5 g of Select Agar (Invitrogen, California, USA) were added to 440 ml of dist. $H_2O$ and autoclaved. The water agar was kept liquid at 50 °C in a water bath and mixed with the 50 ml of M9 salts (10x concentrated), 500 µl of 1 M $MgSO_4$ and 5 ml of 20 % citrate. The volume was adjusted to 500 ml with sterile dist. $H_2O$.

### 3.2.4 Additional chemical compounds

Depending on the bacterial strains manipulated and the different requirements due to the diversity of the experiments, several compounds were added to the media, LB or minimal medium. They are listed in Table 3-3. For each compound, the solvent, the concentration of the stock solution and the final concentration in the medium are detailed.

**Table 3-3: Antibiotics and other chemical compounds added to the media**

| Compound | Solvent | Stock solution | Final concentration |
|---|---|---|---|
| 5-Fluoroorotic acid | 0.5 M Tris-HCl (pH 8.0) | 30 mg/ml | 300 µg/ml |
| Chloramphenicol | $H_2O$ | 15 mg/ml | 15 µg/ml |
| Gentamycin | $H_2O$ | 30 mg/ml | 30 µg/ml |
| Kanamycin | $H_2O$ | 50 mg/ml | 50 µg/ml |
| Nalidixic acid | $H_2O$ | 8 mg/ml | 8 µg/ml |
| Piperacillin | $H_2O$ | 40 mg/ml | 40 µg/ml |
| Potassium tellurite | $H_2O$ | 40 mg/ml | 40 µg/ml |
| Tetracycline | MetOH | 15 mg/ml | 15 µg/ml |
| Uracil | $H_2O$ | 2 mg/ml | 20 µg/ml |

## 3.3 DNA work

### 3.3.1 Enzymatic modification of nucleic acids

#### 3.3.1.1 DNA digestion using restriction endonucleases

Restriction endonucleases were provided by two main companies: Fermentas (Germany) and New England Biolabs, NEB, (Frankfurt am Main, Germany). The cleavage of a DNA molecule was generally done for one hour at 37 °C. When required, the time and temperatures were modified accordingly to the suppliers' protocol (FastDigest® enzyme, Fermentas and Time-Saver™, NEB). The volume of the reaction was set up in a range from 20 to 100 µl mixing always a purified DNA solution (~ 1 µg) with restriction buffer (10x concentrated) and bovine serum albumin (BSA, 10 mg/ml) when required and the desired enzyme(s) (generally 1 unit).

#### 3.3.1.2 Assembly of DNA fragments via ligation reaction

If a single restriction enzyme was used to cleave the cloning vector and the DNA fragment to insert, the digested vector was initially treated with an alkaline phosphatase (1 U/µl) (Shrimp alkaline phosphatase, SAP, Fermentas). The phosphatase is designed to remove the 5'-phosphate and prevent the vector from self-ligation before inserting the DNA fragment. The reaction was generally carried out by adding 1 U of SAP per µg of linear DNA in 20 µl of 1x buffer (provided as 10x concentrated). The phosphatase was inhibited by incubation of the reaction mixture for 15 min at 65 °C.

The T4 DNA Ligase (5 U/µl) was provided by Fermentas. The reaction was carried out in a final volume ranging between 10 and 20 µl and incubated overnight at 16 °C. The molar ratio of 3:1 was

used for the ligation of DNA insert into a cloning vector, respectively. The enzyme was inactivated at the end of the process by incubation of the reaction mixture at 65 °C for 10 min.

### 3.3.2 Design and synthesis of primers

For Polymerase Chain Reaction (PCR) and Arbitrary Primed-PCR (AP-PCR) as well as for sequencing, primers were designed with the Primer3-web (v. 0.4.0) interface of the Primer3 software (Rozen and Skaletsky, 2000). Oligonucleotides were ordered via Eurofins MWG Operon (Ebersberg, Germany) as salt free with a synthesis scale of 0.01 µmol. All the primers used for the PhD work are summarized in Tables 8-1, 8-2, 8-3 and 8-4.

### 3.3.3 Polymerase Chain Reaction (PCR)

The general settings for the PCR amplification were based on the general program given in Table 3-4. The following reaction mixture is shown as a general example:

| | | | |
|---|---|---|---|
| DNA template | x | µl | |
| Buffer $[10x\ conc.]_i$ | 5 | µl | $[1x\ conc.]_f$ |
| dNTPs $[10\ mM]_i$ | 1 | µl | $[0.2\ mM]_f$ |
| Primer A $[10\ \mu M]_i$ | 1 | µl | $[0.2\ \mu M]_f$ |
| Primer B $[10\ \mu M]_i$ | 1 | µl | $[0.2\ \mu M]_f$ |
| Taq DNA polymerase [5 U/µl] | 0.3 | µl | |

All the components were mixed with distillated water to reach 50 ml. The concentrations are introduced between square brackets and the indices $[\ ]_i$ and $[\ ]_f$ designate the initial and final concentrations, respectively. Primers A and B correspond to the ends of the fragment to be amplified by PCR. The buffer and Taq polymerase were provided with the Taq DNA Polymerase kit by QIAGEN (Hilden, Germany). The DNA template is described here as x µl. The amount varied in function of the type of DNA (plasmid, gDNA, bacterial colony), its concentration after extraction and purification and the aim of the PCR.

The concentrations were estimated using the NanoDrop1000 (Thermo Fisher Scientific, Detroit, USA). The amount of DNA template used for the PCR ranged from 50 to 500 ng. PCR reactions were conducted in a Mastercycler ep Gradient (Eppendorf, Hamburg, Germany).

For the purpose of a colony-PCR the biomass was gathered with a sterile toothpick and mixed with the PCR components. In this case, only the initial denaturation step was extended until 5 min instead of 2 min in order to break the cells and release the DNA for further PCR steps.

**Table 3-4: Description of a general PCR program**

| PCR steps | | |
|---|---|---|
| Initial denaturation | 95°C | 2 min |
| Denaturation | 95°C | 25-45 sec |
| Annealing | 56-62°C | 25-45 sec |
| Extension | 72°C | 55-90 sec |
| Final extension | 72°C | 2-5 min |
| **30 cycles of {Denaturation-Annealing-Extension}** | | |

### 3.3.4 DNA Gel Electrophoresis

Charged DNA molecules were separated and visualized by electrophoresis. Depending on the experiments, agarose gels were loaded with different types of DNA: plasmids, linearized vectors, PCR products, digested genomic DNA. Due to the different sizes of these molecules several percentage of agarose gels were used during the PhD work: 0.6 % (w/v) for large molecules of DNA (e.g. genomic DNA), 0.8 to 1 % (w/v) for molecules ranging in size between 800 bp and 6 kb (PCR products or plasmids) and 1.2 % (w/v) for smaller molecules between 50 and 750 bp (e.g. small PCR products). The agarose powder was dissolved in TAE buffer (1x), diluted from a 50x concentrated stock, see the composition below.

For 1 l of 50x concentrated TAE buffer, the following chemical compounds were used as described hereafter:

| | | |
|---|---|---|
| Tris | 242 | g |
| Glacial acetic acid | 57.1 | ml |
| EDTA (0.5 M, pH 8.0) | 100 | ml |

All the chemical compounds were mixed with distillated water to reach 1 l and further diluted to reach the desired concentration. The whole equipment was provided by Biometry (Göttingen, Germany).

Accordingly to the expected size of the DNA and to the different agarose gels employed, different DNA rulers were used. For smaller fragments the GeneRuler™ 50 bp DNA ladder (50 to 1 000 bp) was loaded on the gel. For middle sized and larger fragments the GeneRuler™ 1 kb DNA ladder was used (Fermentas).

The gels were run with a constant voltage which varied between 80 and 100 V depending of the size of the DNA fragments. At the end of the migration the gels were stained in an ethidium bromide

solution. After washing the gels in distillated water, the separated fragments were visualized with UV light. Pictures of the gels were taken with the Gel Documentation system from Bio-Rad (Life Science Group, München, Germany) to be computationally analyzed.

### 3.3.5 Microbial transformation with plasmid DNA

#### 3.3.5.1 Preparation of *Escherichia coli* competent cells for chemical competence

In order to transform *E. coli* strains (DH5α and CC118λ*pir*) chemically with plasmid DNA, the cells were prepared for competence following the protocol from PROMEGA with some minor modifications. A single colony was incubated overnight in Φ Broth medium (see section 3.2.2) at 37 °C under shaking (180 rpm). Four milliliters of the culture were added to 100 ml of Φ Broth medium and incubated for 2 to 3 h at 37 °C under shaking until the culture reached an optical density of $OD_{600nm}$= 0.5 to 0.6. The culture (100 ml) was harvested by centrifugation at 7,000 rpm for 15 min at 4 °C. The supernatant was discarded and the pellets re-suspended with 25 mL of ice-cold TFB1 buffer (see Table 3-5). After 45 min incubation on ice the cells were centrifuged a second time at 7,000 rpm for 15 min at 4 °C. The supernatant was discarded and the pellets re-suspended with 2.5 ml of ice-cold TFB2 buffer (see Table 3-5). Aliquots consisting of 100 µl of the competent cells were kept at -80 °C.

**Table 3-5: Composition of both TFB buffers**

| **TFB1 buffer** | | **TFB2 Buffer** | |
|---|---|---|---|
| $CH_3CO_2K$ | 30 mM | MOPS (pH 6.5) | 10 mM |
| $CaCl_2$ | 10 mM | $CaCl_2$ | 75 mM |
| $MnCl_2$ | 50 mM | $RbCl_2$ | 10 mM |
| $RbCl_2$ | 100 mM | glycerol | 15 % (v/v) |
| glycerol | 15 % (v/v) | | |

The final concentrations are given in Table 3-5. The pH of the TFB1 buffer was adjusted to 5.8 with 1 M of acetic acid. The pH of the TFB2 buffer was adjusted to 6.5 with 1 M of potassium hydroxide. Both solutions were filter-sterilized (0.2 µm) and stored at 4 °C.

#### 3.3.5.2 Chemical transformation of *E. coli* strains

Chemocompetent *E. coli* cells (100 µl) were incubated with 500 ng of plasmid DNA stored on ice for 30 min before conducting the heat shock at 42 °C for 1 min. The mixture was then allowed to stay 2 min on ice before adding 400 µl of liquid LB medium and shaking (180 rpm) at 37 °C for 1.5 h. Part of the culture (50, 100 and 200 µl) was further plated on LB agar supplemented with the required antibiotics and chemicals in order to select the transformed cells. The commercially available pCR®

2.1-TOPO® and pGEM®-T Easy vectors, used for the cloning of intermediate PCR products, were transformed into *E. coli* DH5α. For the establishment of the mini-Tn*5* derivatives, pBAM1 and pJMT6 were transformed into *E. coli* CC118λ*pir* (see Table 3-1).

#### 3.3.5.3 Preparation of *Pseudomonas putida* competent cells for electroporation

The preparation of *P. putida* cells for electroporation followed a protocol adapted from (Holtwick *et al.*, 2001). A single colony was inoculated in 10 ml LB medium and incubated at 30 °C overnight under continuous shaking (180 rpm). Cells were collected by centrifugation at 7,000 rpm for 10 min at 4 °C. The supernatant was discarded and the pellet was washed twice with 10 ml of ice-cold washing buffer (10 % (v/v) glycerol). After the last centrifugation step the supernatant was discarded and the pellet was re-suspended with 100 µl of ice-cold washing buffer. The competent cells were kept on ice for immediate use.

#### 3.3.5.4 Electroporation of *P. putida* strains

Fresh electrocompetent cells (100 µl) were mixed with 500 ng of plasmid preparation. Fifty microliters of the mixture were transferred to a 2-mm-electroporation cuvette (PEQLAB Biotechnology GmbH, Erlangen, Germany). The following settings were applied to the electroporator (Gene Pulser® from Bio-Rad): 2.4 kV (voltage), 25 µF (capacitance) and 600 Ω (resistance). After electroporation 450 µl of LB medium were added and the cells were incubated at 30 °C for 1.5 h under continuous shaking (180 rpm). Parts of the culture (50, 100 and 200 µl) were plated on LB supplemented with uracil and the appropriated antibiotics.

#### 3.3.5.5 Bacterial conjugation: transfer of plasmids to *P. putida* via triparental mating

Both mini-Tn*5* derivatives carrying plasmids, described hereafter, were transformed into *E. coli* CC118λ*pir* due to the presence of the π protein needed for the plasmids to replicate. The transfer of the DNA vectors from *E. coli* to the *P. putida* chromosome necessitated the help of a third strain, *E. coli* HB101 (see Table 3-1). This helper strain carries the pRK600 plasmid encoding the *tra* and *mob* genes, from RK2 (RP4) plasmid (Thomas and Smith, 1987), necessary for the mobilization and transfer of the desired plasmid. The three bacterial strains were inoculated overnight at 30 °C and 37 °C for *P. putida* and *E. coli* strains, respectively, under continuous shaking (180 rpm). Mobilization was carried out with a mixture of these three strains in a 1:1:1 ratio. The pellets were collected by centrifugation of 2 ml of the cultures at 10,000 rpm for 5 min. Supernatants were discarded and the pellets were washed twice with 1 ml of LB medium. After the last washing step 150 µl of LB were used to re-suspend the pellets and mix them together in a single tube. The mixture was applied to a 0.45 µm filter (Millipore™, Massachusetts, USA) on a LB plate. The filter was then dried under sterile conditions at room temperature for 15 min and incubated at 30 °C for 7 to 8 h. A solution of

10 mM of $MgSO_4$ was prepared and 2 ml were used to detach the bacterial drop from the filter. The suspension (50, 100 and 200 µl) was plated on LB medium supplemented with NA and the appropriate antibiotics.

### 3.3.6 Southern blot hybridization

The uniqueness of the insertion event for each of the mini-Tn*5* derived mini-transposons and the confirmation of their presence in the genome were verified by Southern blot hybridization (O'Toole *et al.*, 1999). Four to five microgram of genomic DNA (from *P. putida* SMT or TMT mutants) were digested with an appropriate restriction enzyme. After analysis of the first 600,000 bp of the *P. putida* KT2440 genome (NCBI, http://www.ncbi.nlm.nih.gov/), restriction enzyme *Eco*RV was chosen for the digestion of the genomic DNA. The site recognized by *Eco*RV was found with an average of 16 cuts per 50,000 bp, providing in average fragments of 3,125 bp. The choice of the *Eco*RV enzyme was confirmed by the CMR website (http://cmr.jcvi.org/cgi-bin/CMR/shared/RestrictionDigestSummary.cgi) which indicated 1,318 cuts in the whole chromosome excising fragments of 4,690 bp on average. This size allowed the fragments to be accurately visualized further in the experiments. Moreover, the enzyme did not cut within the mini-Tn*5 KpF* which allowed the probe to finally attach within the mini-transposon. The restriction mixture was then separated on a 0.8 % agarose gel. In order to analyze the data at the end of the experiments, a picture of the ethidium-bromide stained gel was taken (section 3.3.4). The agarose-embedded fragmented gDNA was further transferred overnight to a positively charged nylon membrane (Amersham Hyperbond N+, GE Healthcare; Freiburg, Germany). The kanamycin and tellurite probes were amplified by PCR from the Km resistance gene (miniTn5 *KpF*) and the *telA* gene of the Tel cassette (miniTn5 *TF*), respectively using the following sets of primers: Km-pBAM F/Km-pBAM R and Tel fw/Tel rv (see Table 8-1). The probes were purified and also used as positive controls on the gels (in general the purified PCR products were diluted 100 to 400 times and 20 ng were loaded on the gel). Labeling of the probes was then carried out using the Amersham ECL Direct Nucleic Acid Labeling and Detection Systems kit (GE Healthcare) as described in the provider's instructions; a polymer-complexed peroxidase and gluteraldehyde were involved. The nylon membrane was placed in a hybridization tube and incubated overnight at 42 °C with hybridization buffer containing the labeled probe in a Hybridiser HB-1D (Techne, Staffordshire, UK). Each probe was used for the labeling of a single membrane. After washing the membrane, the detection reagents were directly mixed and applied to it. Reagent 1 contained hydrogen peroxide which was further reduced by the peroxidase already present in the probe sample and allowed the luminol contained in the reagent 2 to produce light. The chemiluminescence was then detected with the Luminescent Image Analyzer LAS-3000 from Fujifilm (Tokyo, Japan). In this way homologous sequences in the genomic DNA fragments were revealed. Samples with a single band appearing on the membrane were kept for

further experiments. If more bands appeared, this meant that the mini-Tn*5* derivatives were inserted at different positions in the chromosome and the mutants were removed from the library.

### 3.3.7 Complementation of a knockout gene and its wild-type

The loss of function of a knockout gene was compensated by complementation with the wild-type gene. For this purpose plasmid pBBR1MCS-5, able of replication in *P. putida* strains, was used for cloning of the target gene. Genomic DNA of *P. putida* TEC1 strain was used to amplify the wild-type gene by PCR with specific primers designed to add *Xba*I and *Sac*I restriction sites at the ends. After restriction of both the purified PCR fragment and the vector, ligation was carried out overnight as mentioned previously. The ligation mixture was transformed into *E. coli* DH5α competent cells. *E. coli* cells carrying the vector were selected on LB agar supplemented with Tc. Colonies were harvested and inoculated for plasmid isolation. Digestion of the vector with *Xba*I enzyme confirmed the insertion of the targeted gene into the vector. Electroporation of the modified vector into the deficient cells was selected on the appropriated medium supplemented with Gm.

## 3.4 Tools for the construction and the detection of the mini-Tn*5* derivatives

### 3.4.1 Construction of mini-Tn*5* transposon derivatives

The construction of the two mini-Tn*5* derivatives consisted of the modification of the available mini-Tn*5 Km* (pBAM1 vector; kindly provided by V. de Lorenzo) and mini-Tn*5 Tel* (pJMT6), see Table 3-1, and will be presented in the following paragraphs.

#### 3.4.1.1 pBAM/*KpF*, mini-Tn*5 KpF* carrying vector

In a first step, modifications were brought to the original pBAM1, mini-Tn*5* transposon carrying vectors. The *pyrF* operon was amplified by PCR from the genomic DNA of *P. putida* KT2440 using the following primers: *pyrF*1F and *pyrF*2R flanking the targeted DNA segment with *Bam*HI and *Hin*dIII restriction sites. The corresponding 876 bp fragment was then cloned into *Bam*HI/*Hin*dIII sites of the patt*FRT* vector (De las Heras *et al.*, 2008) downstream the flippase Recognition target (*FRT*) sequence to create the patt*FRT-pyrF* vector. The *FRT* fragment is considered as being a small DNA fragment due to its 48 bp size. The whole segment <*FRT*::*pyrF*> was amplified using the primers: Fp 1F(*Not*I) and Fp 2R (see Table 8-1) which added *Not*I restriction sites at both ends of the PCR fragment. To generate plasmid pBAM1/*KpF* (see Table 3-1) the resulting fragment was used for ligation with the *Not*I-restricted pBAM1 vector. Correct integration and orientation of the fragment downstream of the Km resistance cassette was selected by restriction analyses. Finally the vector was transformed into *E. coli* CC118λ*pir* as described previously in section 3.3.5.2.

### 3.4.1.2 pUT/*TF*, mini-Tn*5* *TF* carrying vector

In order to construct the second mini-Tn*5* derivative, a single *FRT* fragment was amplified with primers F 10F and F 11R from the patt*FRT* vector and cloned into the pGEM®-T Easy vector (Promega, Madison, USA). Restriction of pGEM<*FRT*> and pJMT6 with *Not*I, providing the miniTn*5* *TF* for the Tellurite selection, was followed by overnight ligation using T4 DNA Ligase (Fermentas, Germany). Both modified vectors were chemically transformed and conserved in *E. coli* CC118λ*pir* as described previously in section 3.3.5.2.

## 3.4.2 Arbitrary Primed-PCR - Mapping of the position of the mini-Tn*5*

The position of the mini-Tn*5* transposon derivatives in the chromosome of *P. putida* SMT and TMT strains was mapped using the Arbitrary Primed-PCR protocol. Two pairs of primers were used in two separate rounds of amplification as presented in Figure 3-1.

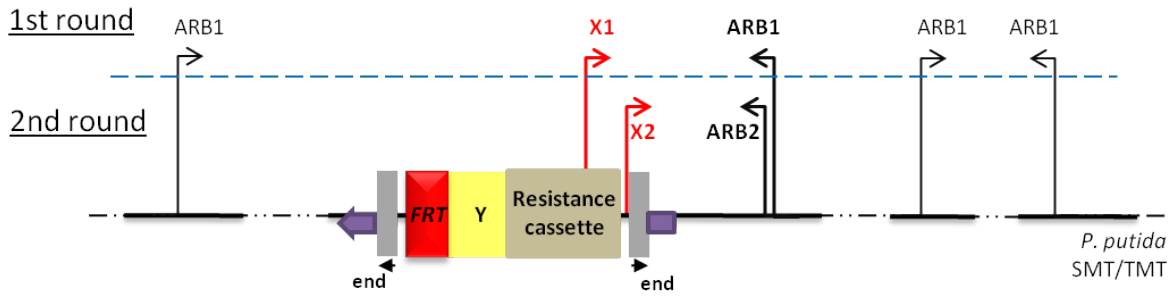



**Figure 3-1: Representation of the Arbitrary Primed-PCR for mapping of a mini-transposon derivative.** The scheme is adapted to correspond to both *P. putida* SMT and TMT cases. The mini-Tn*5* derivative is symbolized by two ends (gray boxes), one FRT fragment (red box), one resistance cassette (tan color) and an eventual third fragment, Y (yellow box). The disrupted gene is represented by a separated lila arrow framing the mini-Tn*5*. Above the dashed blue line the primers participating in the 1st round of AP-PCR are indicated. ARB1 is a degenerated primer and is able to anneal at several places in the chromosome. The primer X1 represents either the APPCR-Km1 for mapping of the mini-Tn*5* *KpF*, or the kilA rev for mapping of the mini-Tn*5* *TF*. The 2nd round is symbolized by the black and red arrows situated below the dashed blue line. ARB2 corresponds to the known sequence of ARB1. X2 represents either APPCR-Km3, if AP-PCR Km1 was used, or F 10F, if kilA rev was used in the first round. The primers in bold (name and arrow) show the two pairs with the highest chances to generate a PCR product small enough to be analyzed.

As can be seen, one specific primer X1 binding inside of the end of the mini-Tn*5* *KpF* or mini-Tn*5* *TF*, and one primer of random sequence ARB1 binding randomly to the chromosome were used for the first round of AP-PCR. All the primers from this paragraph are listed in Table 8-2. The second primer pair was composed of two specific primers. On one side the X2 primer attached to the external part of the mini-transposon derivatives (same direction as X1). On the other side ARB2 attached to the known part of the sequence of ARB1. The composition of the PCR mixture was based on the one described previously (section 3.3.3) but with minor modifications. Due to the higher GC content of

the genomic DNA (61.5 %), the mixture was supplemented with $MgCl_2$ and dimethyl sulfoxide (DMSO), as shown below:

| | | | |
|---|---|---|---|
| DNA template | x | µl | |
| Buffer $[10x\ conc.]_i$ | 5 | µl | $[1x\ conc.]_f$ |
| dNTPs $[10\ mM]_i$ | 1 | µl | $[0.2\ mM]_f$ |
| ARB1 primer | 1 | µl | $[0.2\ \mu M]_f$ |
| Primer X1 $[10\ \mu M]_i$ | 1 | µl | $[0.2\ \mu M]_f$ |
| DMSO $[\geq 99.9\ \%]_i$ | 1 | µl | $[\geq 2.0\ \%]_f$ |
| $MgCl_2$ $[25\ mM]_i$ | 6 | µl | $[3\ mM]_f$ |
| Taq DNA polymerase [5 U/µl] | 0.3 | µl | |

The PCR reaction was carried out in a final volume of 50 µl. The volume of DNA template added to the mixture was dependent on the determined concentration. The $MgCl_2$ was supplied by the Taq DNA Polymerase kit and the DMSO was bought from Sigma-Aldrich (Missouri, USA). Primer X1 corresponded either to APPCR-km1 or kilA rev for amplification with the mini-Tn*5 KpF* or mini-Tn*5 TF*, respectively. The steps of the 1st round of AP-PCR were composed of the cycles described in Table 3-6.

**Table 3-6: Description of the 1st round of the AP-PCR program**

| **AP-PCR steps (1st round)** | | |
|---|---|---|
| Initial denaturation | 95 °C | 5 min |
| Denaturation | 95 °C | 30 sec |
| Annealing | 30 °C | 60 sec |
| Extension | 72 °C | 90 sec |
| **6 cycles of {Denaturation-Annealing-Extension}** | | |
| Denaturation | 95 °C | 30 sec |
| Annealing | 52 °C | 30 sec |
| Extension | 72 °C | 90 sec |
| **30 cycles of {Denaturation-Annealing-Extension}** | | |
| **Final extension** | **72°C** | **4 min** |

From this reaction 2 µl were used as DNA template for the second round of amplification in a final volume of 100 µl as explained below.

| | | | |
|---|---|---|---|
| DNA template | 2 | µl | |
| Buffer [10x conc.]$_i$ | 10 | µl | [1x conc.]$_f$ |
| dNTPs [10 mM]$_i$ | 2 | µl | [0.2 mM]$_f$ |
| ARB2 primer | 2 | µl | [0.2 µM]$_f$ |
| Primer X2 [10 µM]$_i$ | 2 | µl | [0.2 µM]$_f$ |
| DMSO [≥ 99.9 %]$_i$ | 2 | µl | [≥ 2.0 %]$_f$ |
| $MgCl_2$ [25 mM]$_i$ | 12 | µl | [3 mM]$_f$ |
| dist. $H_2O$ | 67.5 | µl | |
| Taq DNA polymerase [5 U/µl] | 0.5 | µl | |

The program specifically used for the second round of AP-PCR is described in Table 3-7.

**Table 3-7: Description of the 2nd round of the AP-PCR program**

| **AP-PCR steps (2nd round)** | | |
|---|---|---|
| Initial denaturation | 95 °C | 1 min |
| Denaturation | 95 °C | 30 sec |
| Annealing | 56 °C | 30 sec |
| Extension | 72 °C | 90 sec |
| Final extension | 72 °C | 4 min |
| **30 cycles of {Denaturation-Annealing-Extension}** | | |

PCR products were cleaned up using the QIAquick PCR Purification kit (QIAGEN) and sent for sequencing (Eurofins MWG Operon) with the external primer X1. Results were analyzed and compared by blasting with the genome of *P. putida* KT2440 (section 3.4.4.1).

## 3.4.3 Sequencing of DNA fragments

### 3.4.3.1 Confirmation of ligations

In order to confirm the insertion of a gene or fragment of interest into the desired vector, restriction analyses were carried out as explained previously (section 3.3.1.1). However, to confirm the absence of base pair substitutions in the inserted DNA fragments, the samples were sent for sequencing. The plasmids were extracted and purified using kits (PROMEGA and QIAGEN) and sent with the appropriate primers for sequencing to Eurofins MWG Operon using the Value Read Tube service.

#### 3.4.3.2 Mapping of the mini-Tn*5* derivatives

For detecting the position of the mini-Tn*5* derivatives in various *P. putida* strains, the extracted genomic DNA was tested with AP-PCR and sequenced at MWG as mentioned previously.

### 3.4.4 Bioinformatics tools

#### 3.4.4.1 DNA sequence analysis tool

The sequencing results were analyzed using different software tools. The Ape software (http://www.biology.utah.edu/jorgensen/wayned/ape/) and the Sequence Scanner v 1.0 software from Applied Biosystems™ were used to read the sequence data. BLAST analyses were carried out applying the Basic Local Alignment Search Tool for nucleotides (blastn) from NCBI (Altschul *et al.*, 1997).

During the mapping of the mini-Tn*5* derivatives, the *P. putida* KT2440 database was directly chosen as reference.

#### 3.4.4.2 Analyses of disrupted and knock-out genes

The data were searched and reported mainly from two websites: Kyoto Encyclopedia of Genes and Genomes (KEGG: http://www.genome.jp/kegg/) (Kanehisa *et al.*, 2006) and the Comprehensive Microbial Resources of the J. Craig Venter Institute (http://cmr.jcvi.org/). Nevertheless in order to obtain the most accurate information the following databases were also consulted: NCBI and the Pseudomonas Genome Database V2 (http://www.pseudomonas.com/).

#### 3.4.4.3 Visualization tool

In order to visualize the position of the different mini-Tn*5* derivatives in the *P. putida* SMT and TMT mutants, the GCView Server (http://stothard.afns.ualberta.ca/cgview_server/) was used. The ‘sequence file’ uploaded corresponded to the one of *P. putida* KT2440 (AE015451.gbk) previously downloaded from NCBI. The ‘gene file’ requested by the server was first downloaded from the NCBI database (AE015451.gff) and modified in order to erase from the Excel file the genes which were not hit by one of the mini-Tn*5* derivatives. In this file each gene, present on the map, was described from the GenBank (NCBI) sequence database, reporting information about the coding sequence, the coding strand and the start and end coordinates. After uploading these two files, the server created a map symbolizing a circular chromosome and showing, upon request of the user, the following features: locus ID of the targeted genes and intergenic regions, GC content and GC skew.

#### 3.4.4.4 *In silico* verification of the Southern blot results

Southern blot hybridization was carried out on *P. putida* SMT and TMT mutants in order to determine the occurrence of the mini-transposon insertion. The signals obtained on the blot were verified by *in silico* digestion of the *P. putida* KT2440 chromosome with *Eco*RV enzyme. The restriction was operated after reporting a window of 15,000 to 20,000 bp framing the hit to the Ape plasmid editor. The length was then estimated between two *Eco*RV sites. The corresponding size was compared with the signal obtained on the blot for the same mutant.

## 3.5 Generation of single- and double-deletion mutants

The general procedure for random genomic deletion in *P. putida* strains leading to the generation of *P. putida* $\Delta_1$ mutants is summarized in three main points hereafter.

**Triparental mating.** The pBBFLP plasmid (Table 3-1) was introduced into chosen *P. putida* TMT cells in order to prepare them for the deletion step itself. The procedure involved different TMT strains, the helper strain *E. coli* HB101 and the *E. coli* DH5α carrying the pBBFLP vector.

**Selection of putative deleted mutants.** Final isolation of the exconjugants was done on the desired medium supplemented with the appropriate antibiotics and further chemical compounds, as specified in the text. A single colony was picked after the deletion step, representing a single potential mutant, and applied on each selective medium needed for the selection of the successfully deleted strain.

**Verification of *bona fide* mutants.** Each exconjugant presenting the expected phenotype were analyzed by AP-PCR and further by sequencing. The same AP-PCR method as mentioned in section 3.4.2 was applied to the selected putative mutants to extract the sequence of the disrupted genes flanking the deleted fragments. Once the genes were obtained or confirmed by sequencing, the genomic rearrangement post-deletion was confirmed by another PCR amplification with specific primers and sequencing of the PCR product.

When the method was applied on single-deleted mutants for the generation of several deletions in the same strain the procedure was modified. Mini-Tn*5 KpF* was inserted in *P. putida* $\Delta_1$-91 mutants, with different colonies pooled together and exconjugants were selected on LB medium supplemented with cit, NA and Km. Colonies (putative *P. putida* $\Delta_1$ SMT mutants) were directly re-suspended from the LB Agar with LB broth containing cit and Km and used as acceptor strain to participate in the next conjugation step aiming at the insertion of mini-Tn*5 TF*. Putative *P. putida* $\Delta_1$ TMT mutants were selected on LB agar supplemented with cit, Tel, Km and NA. The corresponding LB broth (3 ml) was used to re-suspend directly the colonies from the plate and diluted in a total volume of 10 ml of medium. After overnight incubation, part of the culture was involved as acceptor strain in the conjugation with *E. coli* HB101 and *E. coli* CC118λ*pir* pBBFLP. Replication of the FLP carrying

plasmid was selected on medium containing cit, ura, NA and Tc. After overnight incubation allowing the expression of the flippase, colonies were gathered together by adding LB medium, supplemented with ura, cit, NA and Tc and re-suspending the mixture in 50 ml of the same medium. The cells were allowed to grow for 3 to 4 hours and were then spread on medium containing sucrose. After incubation for two to three days at 30 °C colonies were picked and tested on five different LB media all supplemented with cit and ura and containing Km, Tel, FOA or Pip. Due to the absence of intermediary steps and control between each mini-Tn*5* insertion, the piperacillin incorporating plate was used to verify that both mini-Tn*5* were inserted into the genome of *P. putida* $\Delta_1$ TMT and not the whole plasmids. Mutants with the expected phenotypes were then used for a new round of deletion. The strains resulting from genomic excisions were named *P. putida* $\Delta_x$ where x represented the number of successive deletions.

## 3.6 Screening of the different mutants

### 3.6.1 Oxidative stress conditions

The response of the different deleted *P. putida* $\Delta_x$ mutants to oxidative stress conditions was assessed by incubating the strains in medium containing different concentrations of hydrogen peroxide ($H_2O_2$). LB agar plates were prepared with uracil and citrate as carbon source and with $H_2O_2$ with a concentration ranging between 2.and 250 mM. The hydrogen peroxide stock was prepared with a concentration of 1 M by adding 5.6 ml of 30 % $H_2O_2$ into 44.5 ml of water. The solution was filter-sterilized and directly mixed with 50 ml of LB agar. Dilutions were pursued by pouring two plates and mixing the rest of the solution again with 50 ml of LB agar. In total nine different concentrations of $H_2O_2$ were tested and compared with plates containing no hydrogen peroxide. Strains were grown overnight in appropriated medium and one drop (10 µl) was applied for each strain on each agar plate.

### 3.6.2 Bioscreen C MBR as a tool for comparison of bacterial growth

To compare the growth of the different mutants with each other and with the wild-type, growth curve measurements were carried out with a Bioscreen C MBR (Oy Growth Curves Ab Ltd, Helsinki, Finland). In total six strains were cultivated under 18 different growth conditions: the *P. putida* TEC1 wild-type strain, the $\Delta_x$-91 mutant series (*P. putida* $\Delta_1$-91 and *P. putida* $\Delta_2$-91 strain) and the $\Delta_x$-407 mutant series (*P. putida* $\Delta_1$-407 and two different *P. putida* $\Delta_2$-407 strains). These strains were separately inoculated overnight in Falcon tubes in 10 ml of LB medium supplemented with uracil in order to provide the same growth conditions to all of them prior to the beginning of the growth experiments. In total five different 100-well honeycomb format plates (Bioscreen) were filled with M9 minimal medium and uracil supplemented with different carbon and amino acid sources. The general organization of a plate is described in Figure 8-1.

Each carbon source was added to the plates to a final concentration of 10 mM. The concentrations used for L-valine, L-isoleucine and for L-arginine were taken from the literature (Marshall and Sokatch, 1972; Tricot *et al.*, 1991). It was found that growth of *Pseudomonas putida* on valine was stimulated by supplementing the medium with traces of isoleucine (Marshall and Sokatch, 1972). For each strain the growth curve was established in duplicate. Six blanks referred to the six M9 minimal media supplemented with uracil and with either citrate, glucose, succinate, glycerol, benzoate or no carbon source. Each sample was diluted from the overnight culture (taken in middle or late exponential phase) in a pre-defined well to start with an OD = 0.05. The following settings were applied on the Bioscreen controlling computer:

- 30 °C for the incubation temperature and 31 °C for the lid temperature in order to avoid condensation thereby avoiding wrong optical density measurement
- Medium linear shaking
- OD measurement at 600 nm
- Intervals of 15 minutes between each sampling

# 4. Results

## 4.1 Generation of *Pseudomonas putida* single and double mini-transposons mutant libraries

Mini-transposon insertion in the genome of *P. putida* TEC1 allowed the creation of two mutant libraries. The first one corresponded to the transposition of a single mini-Tn*5* into the wild-type cells. The second library was composed of mutants carrying two independent mini-Tn*5* derivatives in their chromosome. These libraries revealed mutants able to grow in LB broth as well as M9 minimal medium supplemented with citrate as sole carbon source. Paragraph will treat separately the generation of these two mutant libraries.

### 4.1.1 Construction of two mini-Tn*5* derivatives for insertion in *P. putida* TEC1

#### 4.1.1.1 The *Saccharomyces cerevisiae* Flp-*FRT* site-specific recombination system

As previously described (sections 3.4.1.1 and 3.4.1.2), the *FRT* fragment was amplified from the patt*FRT* vector to be cloned as a single insert in the mini-Tn*5 TF* or in combination with the *pyrF* operon in the mini-Tn*5 KpF*. The direction of each *FRT* site was decided randomly during the cloning step. However, one condition was absolutely required to assure a potential deletion (as described further in section 4.2.1): one fragment should be oriented towards one end of a mini-Tn*5* and the other one towards the inner part of the second mini-Tn*5*, regardless the derived mini-transposon. As the mini-Tn*5 KpF* was the first derivative created, the orientation of the *FRT* fragment was automatically set up after sequencing and imposed the orientation of the *FRT* for the mini-Tn*5 TF*. Colonies were therefore screened by PCR to obtain the right conditions for the 2$^{nd}$ mini-transposon derivative. The PCR reaction was carried out as described in section 3.3.3 with primers Tel rv and F 10F/F 11R (see Tables 8-1 and 8-3). In the mini-Tn*5 KpF* the *FRT* site was inserted towards the inner part, which corresponds to the Km cassette (Figure 4-1) whereas in the mini-Tn*5 TF* the orientation faced the Oend (Figure 4-2).

#### 4.1.1.2 The pBAM/*KpF* vector: kanamycin resistance construct

Several vectors carrying a Km resistance mediating mini-Tn*5* were available for working with *P. putida* strains (De Lorenzo *et al.*, 1990). Different Km$^R$ mini-transposon carrying vectors were tested for this study (pUT, pBAM1) but the work was finally carried on with the pBAM1 vector (see Table 3-1), itself based on the pUT vector. This vector was designed to become an improved version of the general mini-Tn*5* Km$^R$ carrying plasmid. The pUT/*Km* was modified in order to facilitate the manipulation of the plasmid which was smaller than the previous version. The different features of the pUT/*Km* were either replaced by shorter fragments or completely removed leading to the vector

pBAM1. For this study pBAM1 was modified as described in section 3.4.1.1 to generate pBAM/*KpF* (Figure 4-1) which was used to create different insertions in *P. putida* TEC1 and to generate *P. putida* SMT and further mutant strains. As mentioned in paragraph 2.2.3, an *FRT* fragment corresponds to a directional binding site. Therefore a concrete orientation was given to an *FRT* site by association with a positive or negative sign. Knowing, by sequencing, the orientation of the *FRT* site within the mini-Tn*5* derivative the same sign was also given to the mini-transposon. More details about the sign of the mini-Tn*5* *KpF* (meanwhile the sign of the *FRT* site) and the necessity of associating a direction to it will be explained in section 4.2.1.

The mini-Tn5 *KpF* transposon derivative was the first to be inserted in the chromosome of *P. putida* TEC1 by triparental mating involving *E. coli* CC118λ*pir* pBAM/*KpF* and the *E. coli* HB101 as helper strain. The procedure was applied as described in the paragraph 3.3.5.5.

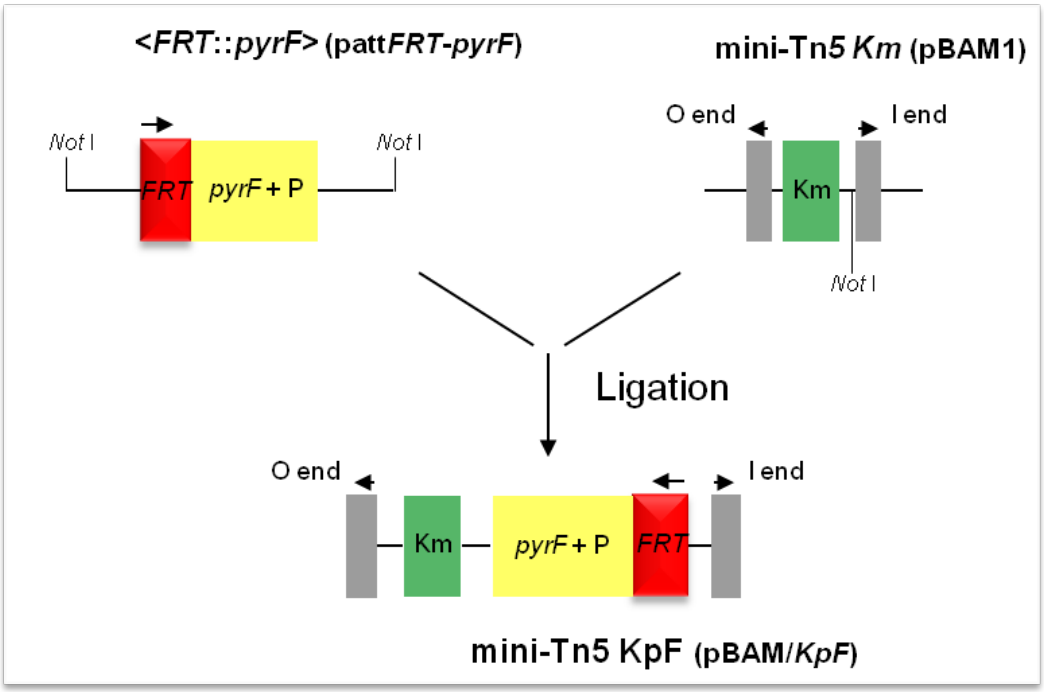



**Figure 4-1: Assembly of the different fragments for the construction of the pBAM/*KpF* vector.**
Based on vector pBAM1 the single *Not*I restriction site was used to insert the <*FRT*::*pyrF*> cassette excised from the patt*FRT-pyrF* vector. The Km cassette was employed as a selection marker to confirm the presence of the mini-Tn*5* derivative in the cells after triparental mating. The *pyrF* operon played a dual role: on the one hand it was possible to use it as second confirmation for the presence of the mini-Tn*5* *KpF*, on the other hand the sensitivity of the strain against 5-FOA could force the homologous recombination between two *FRT* sites. For this purpose the single *FRT* site (red) was inserted towards the inner part of the minitransposon. The yellow boxes marked with "*pyrF* + P" refer to the *pyrF* gene under control of its own promoter. The Km cassette is shown as a green box. The gray rectangles represent each end of the mini-Tn*5*. The orientation of the FRT fragment is shown by a black arrow.

### 4.1.1.3 The pUT/*TF* vector: tellurite resistance construct

Based on pJMT6, the pUT/*TF* was generated following the procedure described in section 3.4.1.2. The main feature inserted in the single *Not*I site of mini-Tn*5 TF* was a single *FRT* fragment without the *pyrF* operon, which was already present in the mini-Tn*5 KpF* (Figure 4-2). The *FRT* site was orientated towards the Oend. Further explanations about the direction of the *FRT* fragment and the mini-Tn*5 TF* itself will be given in section 4.2.1.

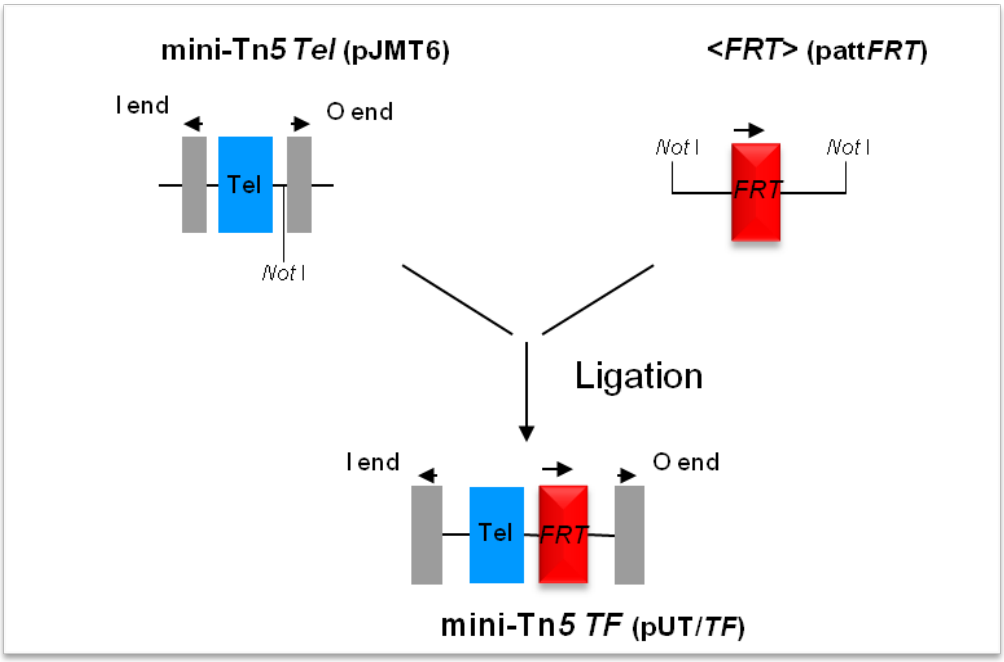



**Figure 4-2: Assembly of the different fragments for the construction of the mini-Tn*5 TF*.**
Based on pJMT6 a single *Not*I restriction site was used to insert a single *FRT* site which was excised from the patt*FRT* vector. The Tel cassette was employed as a selection marker to confirm the presence of the mini-Tn*5* derivative in the cell after triparental mating. The *FRT* fragment (red) was inserted towards the Oend of the minitransposon (symbolized by the black arrow above it). The blue boxes represent the Tel operon and the gray rectangles the ends of the mini-Tn*5*.

*P. putida* mutant strains carrying the mini-Tn*5 TF* were able to grow on medium containing potassium tellurite in comparison with the wild-type strain for which the compound was toxic. Its insertion was easily recognizable due to the production of a characteristic black color by the reduction of the tellurite into tellurium (Figure 4-3).

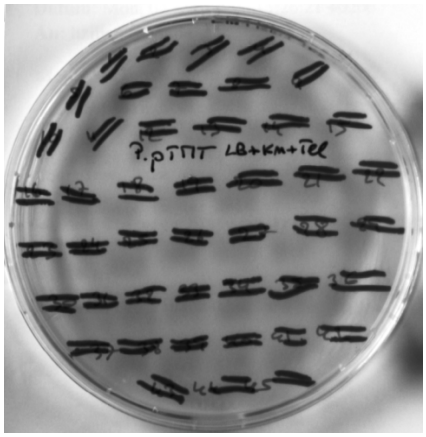



**Figure 4-3: Effect of potassium tellurite on *P. putida* mutant strains**
Different colonies from strains carrying the mini-Tn*5 TF* were streaked on LB plates containing potassium tellurite. Due to the presence of the Tel operon (*kilA*, *telA* and *telB* genes) in the mutant the tellurite was reduced to tellurium producing a characteristic black color.

The pUT/*TF* was used to construct the second mutant library which consisted on *P. putida* TEC1 double mini-transposon mutants (*P. putida* TMT strains).

### 4.1.2 Generation of *Pseudomonas putida* TEC1 mutants carrying single or double mini-transposon insertions

Separated transposition of both mini-Tn*5* *KpF* and *TF* were carried out in *P. putida* TEC1 generating two types of mutants, carrying either a single mini-Tn*5* or two mini-Tn*5* derivatives in their genome. From these experiments, a list of genes non-essential in LB broth and M9 minimal medium with citrate as carbon source was extracted. Furthermore the establishment of a library of double mutants will prepare the final step of genomic deletion and reduction of the genome of a single *P. putida* TEC1 cell.

#### 4.1.2.1 Single mini-transposon mutant library

The first library was based on the insertion of a single mini-transposon derivative in the genome of *P. putida* TEC1 generating Single Mini-Transposon (SMT) strains.

##### 4.1.2.1.1 Establishment of the library

The transfer of pBAM/*KpF* into *P. putida* TEC1 via triparental mating was carried out as described in paragraph 3.3.5.5. The insertion of the mini-Tn*5* *KpF* into the chromosome of *P. putida* TEC1 generated single mini-transposon mutants named *P. putida* SMT strains. Sixty-six potential *P. putida* SMT mutants were selected after plating of the conjugation mixture on M9 minimal medium supplemented with citrate, uracil and kanamycin in order to carry out the first phenotype controls.

All the colonies were streaked on LB medium supplemented with citrate and separately with three different components: Km or Pip or FOA. Sixty-one colonies showed a {$Pip^S$, $FOA^S$ and $Km^R$} phenotype whereas the last five colonies exhibited a {$Pip^R$, $FOA^S$ and $Km^R$} phenotype. These results indicated that in 93 % of the cases true transposition events occurred, which means that only the mini-Tn*5* transposon was transferred to the chromosome and the backbone of the plasmid was lost. This value was slightly lower than the 95 % of efficiency for transposition using the mini-Tn*5* transposon in *P. putida* strain obtained in former experiments (De Lorenzo and Timmis, 1994). This could be explained simply by the fact that only 66 colonies were screened out of one single experiment which did not provide a good significance for the results.

The 61 colonies were further inoculated from the LB + cit + Km plates into M9 minimal medium supplemented with cit and Km overnight at 30 °C. Genomic DNA was also extracted for further experiments from each of the 61 cultures collecting between 600 and 1,500 ng of DNA for the different samples.

##### 4.1.2.1.2 Confirmation of the putative *P. putida* SMT mutants

Out of the 61 extracted gDNA samples, thirty were chosen and used for Southern blot experiment in order to confirm the presence of a single mini-transposon derivative in the genome. The protocol was

applied to 4 µg of the DNA samples as described in paragraph 3.3.6. Pictures of the membranes containing the hybridized DNA were made using the LAS-3000. Each sample revealed a single band with different sizes. As an example, the results of the first 20 *P. putida* SMT used for Southern blot experiments were illustrated in Figure 4-4.

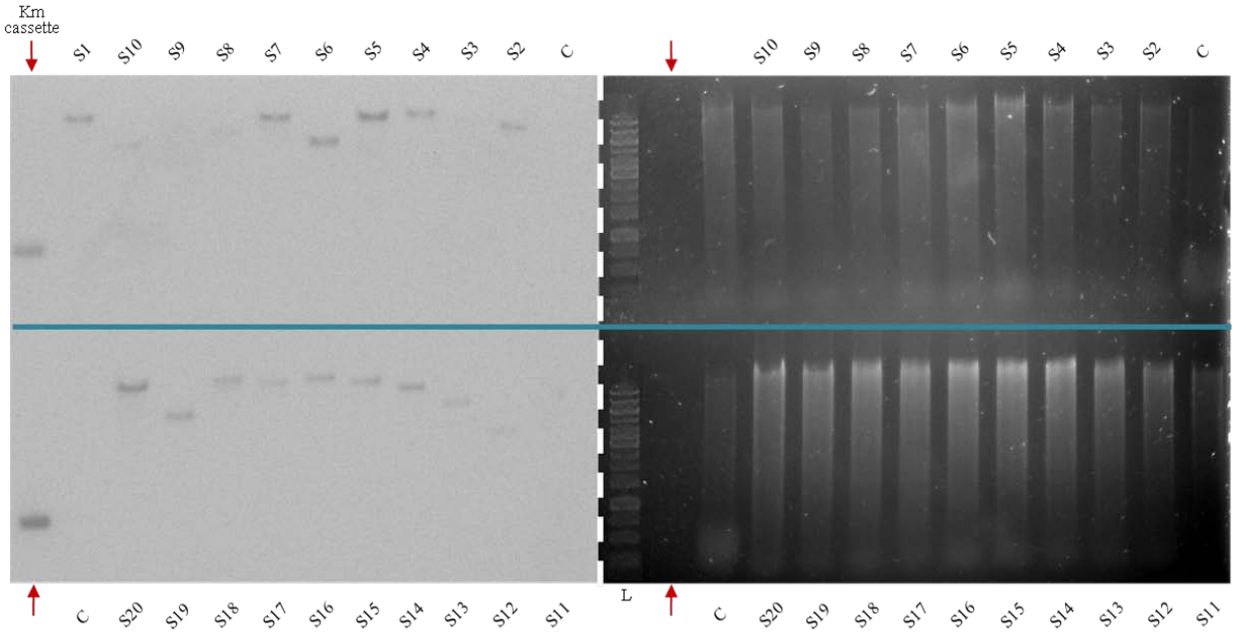



**Figure 4-4: Verification of the presence of the mini-Tn*5 KpF* in the chromosome of *P. putida* SMT strains.** Agarose gels (on the right side) and corresponding blots (on the left side) of the *Eco*RV restricted gDNA of *P. putida* SMT strains and *P. putida* TEC1, used as the negative control (C). The Km probe was used for hybridization and served as positive control (indicated by a red arrow). The pictures above the blue line represent the results obtained with the wild-type strain (C) as well as with the *P. putida* SMT col. S1 to S10. The pictures under the blue line show the results of colony S11 to S20 and the control (C). The numbers S1 to S20 refer to the corresponding *P. putida* SMT colonies. The ladder (L) was applied for the gel electrophoresis but was used as well for the analysis of the blots.

Due to its low concentration the Km probe did not appear on the gels (right side) but was revealed by Chemiluminescence on both blots (left side of Figure 4-4) as a single band corresponding to the 760 bp PCR product. As expected *P. putida* TEC1 (lanes C on the left side of Figure 4-4) did not show any homologous sequence with the Km cassette, which confirmed the use of the strain as a negative control. For the 20 other lanes, corresponding to *P. putida* SMT col. S1 to S20, a single band was visible on the blots. The presence of a unique signal proved that the mini-Tn*5 KpF* was inserted only once into the chromosome of each mutant. Due to the random distribution of the *Eco*RV restriction sites within the chromosome of *P. putida* TEC1, the full digestion with the enzyme led to the release of several fragments of different length. Therefore the presence of two or more mini-Tn*5 KpF* in the genome of a mutant would be detected as two or more signals on the blot for a single mutant.

Furthermore, the blots provided qualitative information regarding the difference in the position of the mini-Tn*5* derivative for each mutant. Each mutant was generated from the same wild-type strain however hybridization of the Km probe with the restricted gDNA of the different colonies led to the exhibition of single bands of different sizes.

As described above, the 20 *P. putida* SMT putative mutants were confirmed by Southern blot experiments to carry a single mini-Tn*5 KpF* in a different position in the genome. The data for the 10 other strains were not presented here but led to the same conclusions concerning a unique insertion at a different place in the genome for each of them. All *P. putida* SMT strains collected after triparental mating were therefore considered as potential mutants and further verified with the mapping of the mini-Tn*5 KpF*.

#### 4.1.2.1.3 Mapping of the mini-Tn*5 KpF* for each SMT mutants

The genomic DNA of each sample was further used for Arbitrary-primed PCR amplification in order to sequence and map the position of the mini-Tn*5 KpF* for the different *P. putida* SMT strains. Using an arbitrary primer in combination with a primer of known sequence (inside the mini-Tn*5 KpF*) in the first round, the whole genome was screened for insertion of the mini-Tn*5* derivative. The second round of PCR was thus used as a normal PCR to amplify and to enrich a specific fragment. Different settings, e.g. annealing temperature or/and time were tried for both rounds of amplification in order to establish the most appropriated protocol. The final combinations were described in paragraph 3.4.2.

The main goal of the AP-PCR was to determine in which gene or intergenic region the mini-Tn*5 KpF* was inserted for each mutant. However, another important feature of the insertion was revealed in the mean time. As previously mentioned in section 4.1.1.2, a positive or negative sign was associated with mini-Tn*5 KpF*, which provided information about the direction of the *FRT* site relatively to the chromosome. For this purpose the APPCR-Km3 primer used for sequencing of the disrupted genomic fragment was designed within the Km cassette towards the Oend of the mini-Tn*5*. The BLAST tool was used to compare the sequencing results with the genome sequence of *P. putida* strain KT2440. The orientation of mini-Tn*5 KpF* and therefore the one of the *FRT* fragment were determined as positive if the query sequence followed the 5’ to 3’ direction of the subject. On the contrary a negative sign was allocated to the mini-transposon when the query sequence matched with the 3’ to 5’ direction of the subject. The orientation of the *FRT* sequence was, thus, easily recognizable during the mapping of the mini-Tn*5 KpF* which was useful for the detection of potential candidate for deletion, as will be explained in section 4.2.1.

In order to establish the mutant library, the 61 putative *P. putida* SMT mutants were undergone to AP-PCR and the position of each of the mini-Tn*5 KpF* was summarized in Table 8-5. In total 51 individual genes were disrupted and three mini-Tn*5* insertions appeared in separated intergenic

regions, meaning that in 88.5 % of the cases the mini-Tn*5 KpF* hit different sequences of the chromosome.

The distribution of the hits was represented in Figure 4-5. The map was created as described in section 3.4.4.3. In the case of intergenic regions hit by the derived mini-transposon, the locus ID (PP) was replaced in the ‘gene file’ by the intergenic region symbol: IR and the locus ID of both the flanking genes. As an example the colony S18 was hit by a mini-Tn*5 KpF* between the genes PP_0780 and PP_0781. In the map as well as in the list of genes and intergenic regions the hit was annotated as IR_0780-81. In order to be represented as an arrow on the map as it was done for the genes, a sign was allocated to the IR. During the establishment of the ‘gene file’ the orientation of the downstream gene was kept for giving an orientation to the IR. Therefore on the map three IRs could be found, two of them on the positive strand and the third one on the negative strand. For IR_0780-81 the legend was positioned on the outer part of the rings due to the positive orientation of the PP_0780 gene.

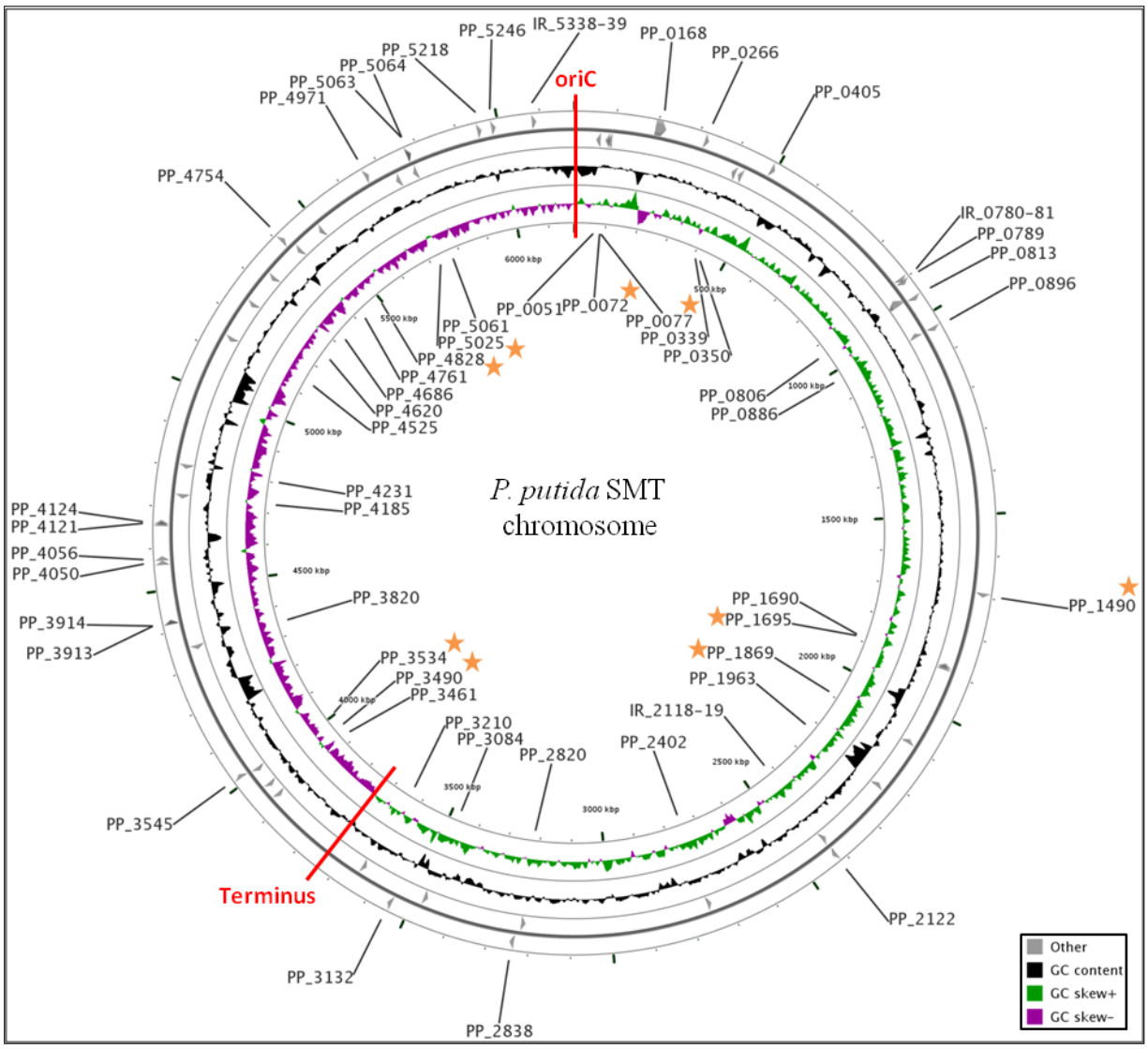



**Figure 4-5: Map of the different positions of the mini-Tn*5 KpF* after insertion in the genome of *P. putida* TEC1.**

Each hit was obtained in separated mutants but combined on the same graphical map for the convenience of the representation. In total 54 independent ORFs and intergenic regions were found after sequencing of the AP-PCR products carried out with the 61 *P. putida* SMT mutant strains. The yellow stars indicate the corresponding *P. putida* SMT mutants selected for the insertion of the mini-Tn*5 TF* transposon. The loci ID (PP for a gene and IR for an intergenic region) are positioned on both sides of the circles at the same coordinate as the corresponding gray arrow indicating the coding strand. GC content (black color) is indicated on the middle ring with values varying above and below the 61.52 % average of the whole chromosome. Fluctuations of the GC skew are reported on the internal ring with positive (green color) and negative (Lila color) values.

The genes IDs (PP) were placed beside the gray arrows following the orientation of the coding strand. The yellow stars indicated the nine mutants selected for the generation of the double mini-transposon mutant library. The strains carrying mutation in the following genes: PP_0072 (*qor-1*), PP_0077 (*betC*), PP_1490, PP_1695, PP_1869, PP_3490, PP_3534, PP_4828 (*cobH*) and PP_5025 (*mdoH*)

were the first to have their mini-Tn*5* mini-transposons mapped and therefore directly used for further experiments.

The mapping by AP-PCR and sequencing confirmed the results obtained after Southern blot experiments, paragraph 4.1.2.1.2. To illustrate this conclusion, three examples of *P. putida* SMT mutants were taken: colonies S6, S12 and S19. For each of them the gene or intergenic region hit by the mini-Tn*5 KpF* was found in Table 8-5 and the corresponding coordinates were extracted from the Pseudomonas Genome Database V2. The method described in section 3.4.4.4 was applied to the three mutants. The results were presented hereafter.

- For col. S6 the mini-Tn*5 KpF* hit the gene PP_1869 precisely at the coordinate: 2 092 827 bp. The genomic sequence from *P. putida* KT2440 was reported to the plasmid editor from 2,080,000 to 1,100,000 bp (Figure 4-5). After *in silico* restriction the fragment containing the mini-Tn*5* gave a size of 6,731 bp. The comparison with the Southern blot obtained after hybridization (Figure 4-4) confirmed these results.
- For col. S12 the mini-Tn*5 KpF* hit an intergenic region situated between the genes PP_2118 and PP_2119, exactly at the coordinate: 2,416,824 bp. The sequence extracted from the NCBI database covered the region framed by the following coordinates: 2,410,000 to 2,426,000 bp. After adding the mini-Tn*5 KpF* into the targeted sequence and *in silico* restriction with *Eco*RV, the size of the fragment framed by two *Eco*RV sites including the mini-Tn*5* derivative corresponded to the size (3,865 bp) of the signal obtained with the blot previously shown in Figure 4-4.
- For col. S19 the mini-Tn*5 KpF* hit the gene PP_3913 precisely at the coordinate: 4,422,076 bp. The sequence between 4,415,000 and 4,435,000 bp was selected for *in silico* reconstruction of the insertion of the mini-transposon. After restriction the fragment was found to have a size of 5,803 bp, corresponding to the band revealed on the blot for this mutant strain.

These comparisons confirmed the presence of a single mini-Tn*5 KpF* in the genome of the different mutants, as visualized after Southern blot hybridization (Figure 4-4). Furthermore, the positions of the mini-transposon in the genome were verified as it was predicted in paragraph 4.1.2.1.2.

#### 4.1.2.1.4 Analysis of the SMT mutants

The sequencing of the AP-PCR products of each of the 54 independent *P. putida* SMT mutants resulted in the first single mutant library highlighting some of the conditionally non-essential genes of *P. putida* TEC1in M9 minimal medium supplemented with uracil and citrate.

In total, three *P. putida* SMT mutants carried a single mini-Tn*5 KpF* within an intergenic region: IR_0780-81, IR_2118-19 and IR_5338-39. The consequence of this genome disruption for each of the

three strains was not confirmed in the laboratory; therefore the cellular roles of the proteins encoded by the framing genes were not taken into account for establishing the following percentages, which meant that the total amount of strains considered was decreased to 51. For a better representation of the different cellular role categories concerned by the disrupted genes, a pie chart was created and shown on. Out of these 51 ORFs four were coding for proteins with unknown function, four for conserved hypothetical proteins and three for hypothetical proteins. In summary 21.5 % of the hits concerned genes whose function has not yet been described. The two main cellular roles represented in the different *P. putida* SMT strains were 'Energy metabolism' and 'Regulatory functions', with eight genes for each category (13.6 %). 'Transport and binding proteins' was the second category with seven different identified genes (11.9 %). 'Cellular processes' and 'Cell envelop' regrouped six and five genes, 10.2 % and 8.5 %, respectively. Three hits (5.1 %) concerned genes coding for proteins participating in a 'Signal transduction' role and four hits represented genes encoding proteins involved in 'Biosynthesis of cofactors, prosthetic groups and carriers' (two hits, 3.4 %) and 'Central intermediary metabolism' (two hits, 3.4 %). The last four categories represented by the disrupted genes and concerning each one single hit (1.7 %) were regrouped into: 'DNA metabolism', 'Fatty acid and phospholipid metabolism', 'Mobile and extrachromosomal element functions' and 'Purines, pyrimidines, nucleosides, and nucleotides'. All the percentages were calculated out of 59 hits, instead of 51, due to the participation of eight encoded proteins to more than one cellular role categories. In total seven proteins played a role in two different categories and one in three categories.

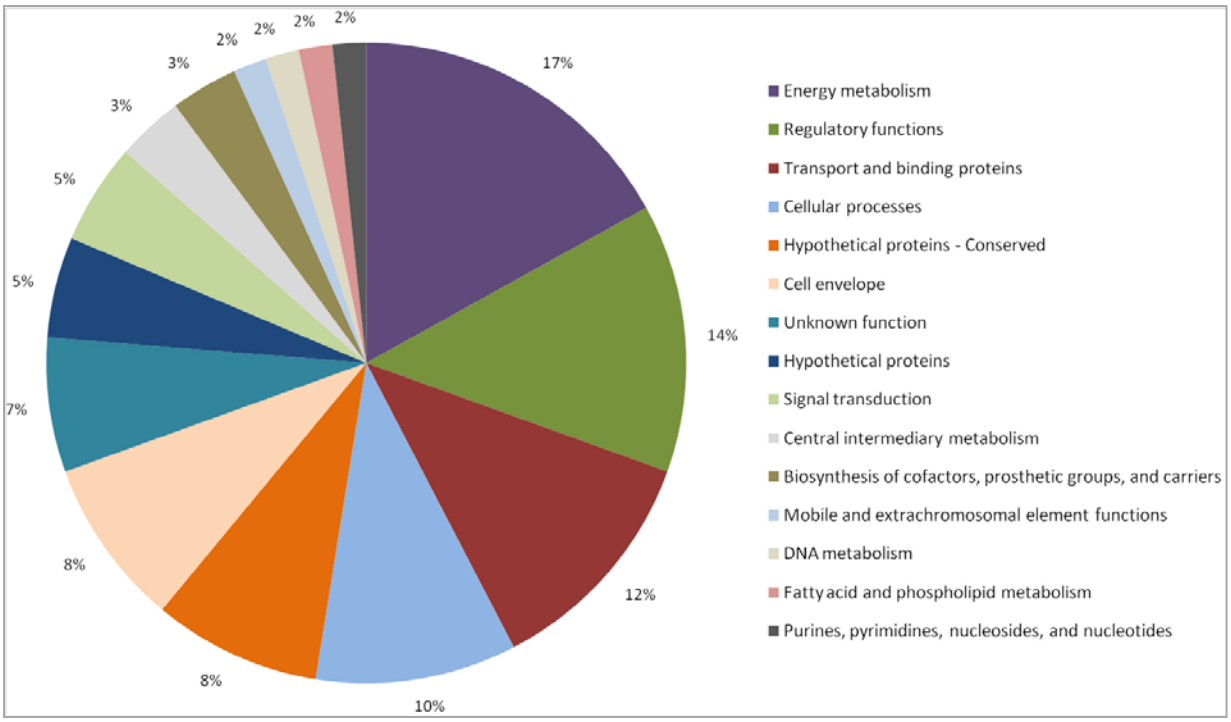



**Figure 4-6: Representation of the different cellular roles concerned by the genes disruption in *P. putida* SMT mutants.**
The pie chart categorizes the disrupted genes of *P. putida* SMT mutant into their different cellular roles. One colored sector represents one category. The percentage of genes composing this category is annotated outside the sector. One cellular role category was not found in the mutant, thus not represented on the pie chart. The percentages are calculated out of 59 strains due to the presence of eight encoded proteins which play a role in two different categories.

Interestingly some of the disrupted genes which play a role in 'cellular processes' and more precisely in cell adhesion such as PP_0168 (*lapA*) (26,049 bp) and PP_0806 (*lapF*) (18,933 bp) have already been described in the genome-wide mutant library previously established and analyzed for the *P. putida* strain KT2440 (Duque *et al.*, 2007). These two ORFs, which are the two longest ORFs in the genome of *P. putida* KT2440, were found to represent genes being mostly hit by mini-Tn*5* transposon which was certainly due to their properties, as will be further discussed in section 5.2.2. Even though in the case of the present study the AP-PCR detected only one hit for each of these ORFs, but this may be due to the low number of analyzed mutants.

As previously announced, 54 independent genes and intergenic regions were hit once by the single mini-Tn*5* *KpF*. Interestingly for some genes (PP_0051, PP_1963, PP_3084, PP_3490 and PP_3534) two individual *P. putida* SMT mutants were found in the single mutant library. Furthermore, the PP_4121 gene (*nuoCD*), subunit of NADH dehydrogenase I, was hit in three different mutant strains.

#### 4.1.2.2 Double mini-transposon mutant library

Nine *P. putida* SMT mutants (SMT 1 to 3 and 5 to 10, Table 8-5) were chosen as previously described for the establishment of the *P. putida* TMT mutant library. The mini-Tn*5 TF* transposon was transferred via triparental mating into the selected mutants, as described in paragraph 3.3.5.5. In this section we described the establishment of the double mutant library and the hypothetical candidates selected for genomic fragment deletions.

##### 4.1.2.2.1 Mapping of the disrupted ORFs

The nine selected *P. putida* SMT strains were pooled together and used as a single recipient cell for the pUT/*TF* delivery. After the triparental mating, single black colonies were randomly picked from the M9 plates and streaked on the different control plates. The black color indicates the presence and expression of the Tel cassette in the cells which supposes that the mini-Tn*5 TF* was inserted into the genome. In total 613 colonies were chosen out of four separated triparental matings involving the same donor, helper and recipient cells.

Out of the 613 chosen colonies, 573 *P. putida* TMT mutants revealed a {$Pip^S$, $FOA^S$, and $(Km, Tel)^R$} phenotype indicating that in 93.5 % of cases true transposition events occurred. The value remained lower than the one estimated by de Lorenzo & Timmis (De Lorenzo and Timmis, 1994); however the same value was found for the *P. putida* SMT mutants (93.0 %, section 4.1.2.1.1) and therefore considered as a new limit for the transposition efficiency in *P. putida* TEC1.

For further analyses, genomic DNA was isolated from all 573 *P. putida* TMT mutants. The DNA concentration ranged from 500 to 1,800 ng/µl. Amounts varying between 0.3 and 0.6 µl of the samples were used as template for AP-PCR amplifications.

Using the AP-PCR and subsequent sequencing analyses, both mini-Tn*5* derivatives were mapped in order to build a library of double mutants. In total more than 1,300 Arbitrary Primed PCR amplifications were carried out in order to describe the position of 789 independent mini-Tn*5* derivatives (*KpF* and *TF*). All 573 *P. putida* TMT mutants were not yet analyzed for the position of both the mini-Tn*5* derivatives simultaneously. The results of the sequencing were summarized in Table 8-7 and simultaneous position of both mini-Tn*5 KpF* and *TF* were located (Table 8-6). For the mini-Tn*5 KpF*, 507 positions were detected and 282 for mini-Tn*5 TF* (a total of 789). These 282 hits were visualized on a graphical map created with the GCView server, as it was previously established for the list of the genes and intergenic regions hit by a mini-Tn*5 KpF* (section 4.1.2.1.3). A single chromosome is symbolized by a series of dark gray rings. For the convenience of the map's interpretation all mini-Tn*5 TF* hits were reported on a single chromosome. The map is presented in Figure 4-7. As previously explained for the creation of the *P. putida* SMT chromosome map

(Figure 4-5) the intergenic regions represented by 'IR_' and the downstream and upstream genes were placed on the same coding strand as the upstream gene.

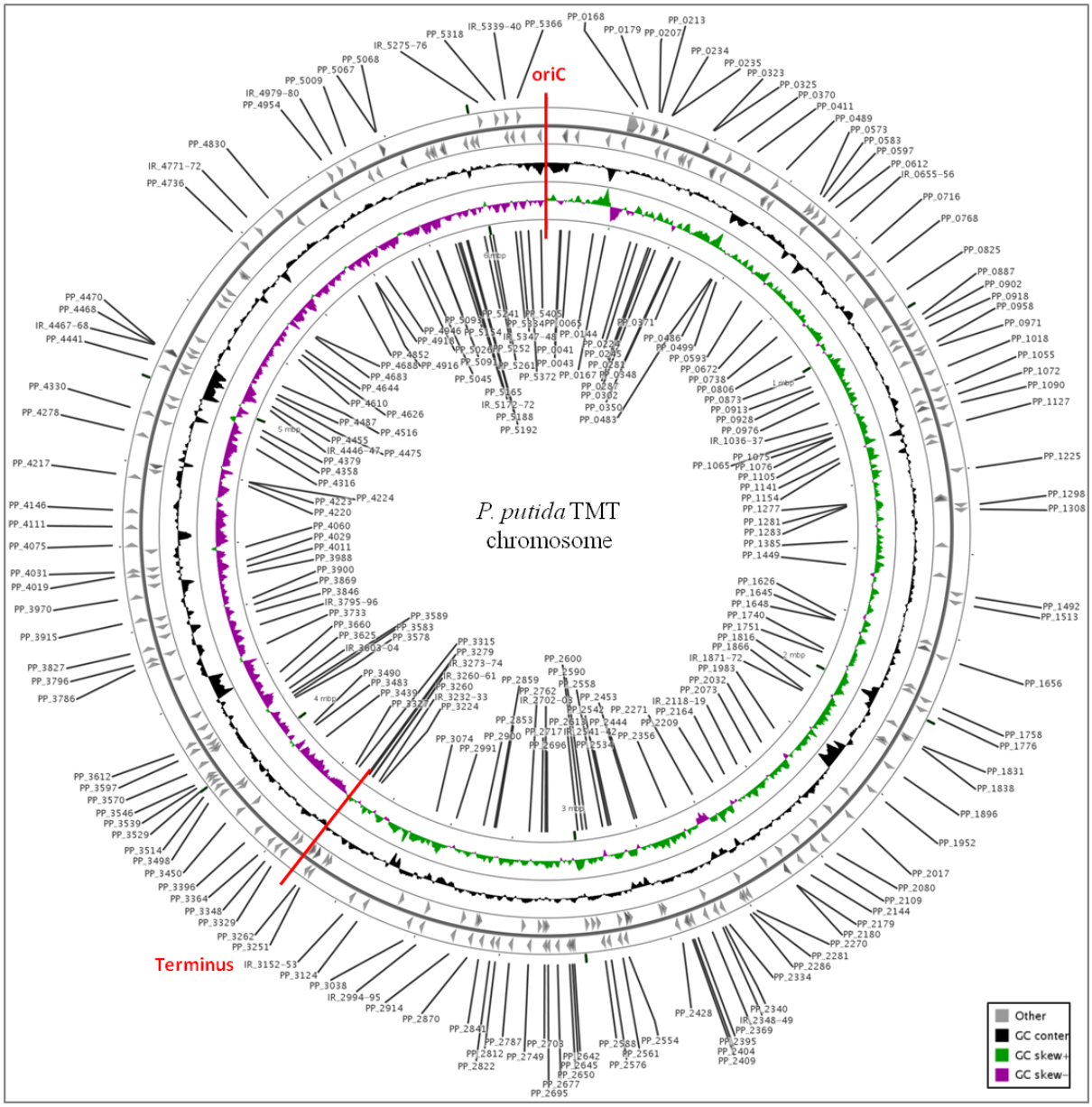



**Figure 4-7: Map of the different positions of the mini-Tn*5 TF* after their insertion in the genome of nine *P. putida* SMT strains.**
Each hit was obtained from different mutants however for the convenience of interpretation the mini-Tn*5 TF* insertions were combined on the same graphical map. In total 255 independent genes and intergenic regions were found after sequencing of the AP-PCR products. The loci ID (PP and IR) are positioned on both sides of the circles. As a central ring, the fluctuations of the GC content are drawn above and below the average percentage of 61.52 %. Variations of the GC skew are represented in green color (positive values) and in lilac color (negative values). From the origin to the terminus of replication (red line), the GC skew is mainly drawn in green. After the Terminus negative values are mostly represented.

Out of the 282 positions extracted from the whole set of *P. putida* TMT mutants, 255 mini-Tn*5 TF* hit the chromosome at a different coordinate. This means that in 90.4 % of the cases the mini-transposon was inserted in a different ORF. In total 18 genes appeared twice or more than twice in the sequencing

results and four mutants were not able to be precisely detected for the mini-Tn*5 TF* insertion due to the hit in a repetitive sequence of the genome. Even though in a particular case (sequencing of mini-Tn*5 TF* for col. T597) the repetitive sequences all belonged to the same gene: PP_0168 (*lapA*) (26,049 bp), the exact position could not be determined. These genes were therefore neither reported on the map nor in the tables.

Out of the 255 independent genes, 22 intergenic regions were hit by a mini-Tn*5* derivative. The final 233 genes were taken into account for the following analyses.

#### 4.1.2.2.2 Analysis of the Two Mini-transposons (TMT) mutants

For each *P. putida* TMT mutant, carrying two mapped mini-Tn*5* derivatives (mini-Tn*5 KpF* and mini-Tn*5 TF*), both disrupted genes were reported in Table 8-6. The quality of the sequencing results was divergent from one sample to the other one. Only nine independent *P. putida* SMT strains were used for the second round of mini-Tn*5* insertion. The sequencing results providing the position of the mini-Tn*5 KpF* in the *P. putida* TMT mutants were expected to match with one of the nine disrupted genes from the initial *P. putida* SMT. Therefore, the position of the mini-Tn*5 KpF* was easier to identify than the one of the mini-Tn*5 TF* even when the resulting sequence was short and the matching with the genome of *P. putida* KT2440 lower than 85 %. The description of the 255 independent genes or intergenic regions hit by the mini-Tn*5 TF* was presented in Table 8-7.

As described in Table 8-7, the loci IDs, genes, product names, K/EC numbers and cellular role categories were given for each gene disrupted by the mini-Tn*5 TF* in the different mutants. When the mini-Tn*5 TF* was inserted between two ORFs, a description of both the flanking genes was provided. However, the influence of the mini-transposon insertion on the transcription of the flanking genes was not assessed; therefore, the cellular role category of each was not taken into account for establishing the following percentages. Some of the encoded proteins participated to two categories; they remained in white color in the Table 8-7.

In order to better understand which disrupted genes were responsible for which cellular roles, a pie chart was created in Figure 4-8. The main defined category represented in Table 8-7 corresponded to ‘Transport and binding proteins’ with 48 hits (18.8 %), spread over six sub-roles: Amino acids, peptides and amines; Anions; Carbohydrates, organic alcohols, and acids; Cations and iron carrying compounds; Porins and Unknown substrate. ‘Regulatory functions’ and ‘Energy metabolism’, with 29 (11.4 %) and 26 hits (10.2 %) respectively, were the most represented categories. Nineteen (7.5 %) and 17 (6.7 %) encoded proteins played a role in the ‘Cellular processes’ and ‘Cell envelope’ categories respectively. ‘Protein fate’, ‘DNA metabolism’, ‘Mobile and extrachromosomal element functions’ and ‘Signal transduction’ were equally represented ~3.1 % of all the categories. Finally ‘Central intermediary metabolism’, ‘Fatty acid and phospholipid metabolism’, ‘Biosynthesis of

cofactors, prosthetic groups, and carriers', Amino acid biosynthesis', 'Protein synthesis' and 'Purines, pyrimidines, nucleosides, and nucleotides' were the less prevalent the groups with only 1 to 7 hits per category. In total 31 proteins were annotated as 'Conserved hypothetical proteins' and 25 others were assigned to 'Unknown functions'. Fourteen *P. putida* TMT mutants carried a mini-Tn*5 TF* insertion in a gene encoding a hypothetical protein not described yet. In total, 26.8 % of the gene hits by the mini-Tn*5* derivative encoded uncharacterized proteins slightly more than the number of hits observed for the mini-Tn*5 KpF*. Since 28 proteins could be assigned to two different cellular role categories, the percentages were calculated out of a total of 261 hits.

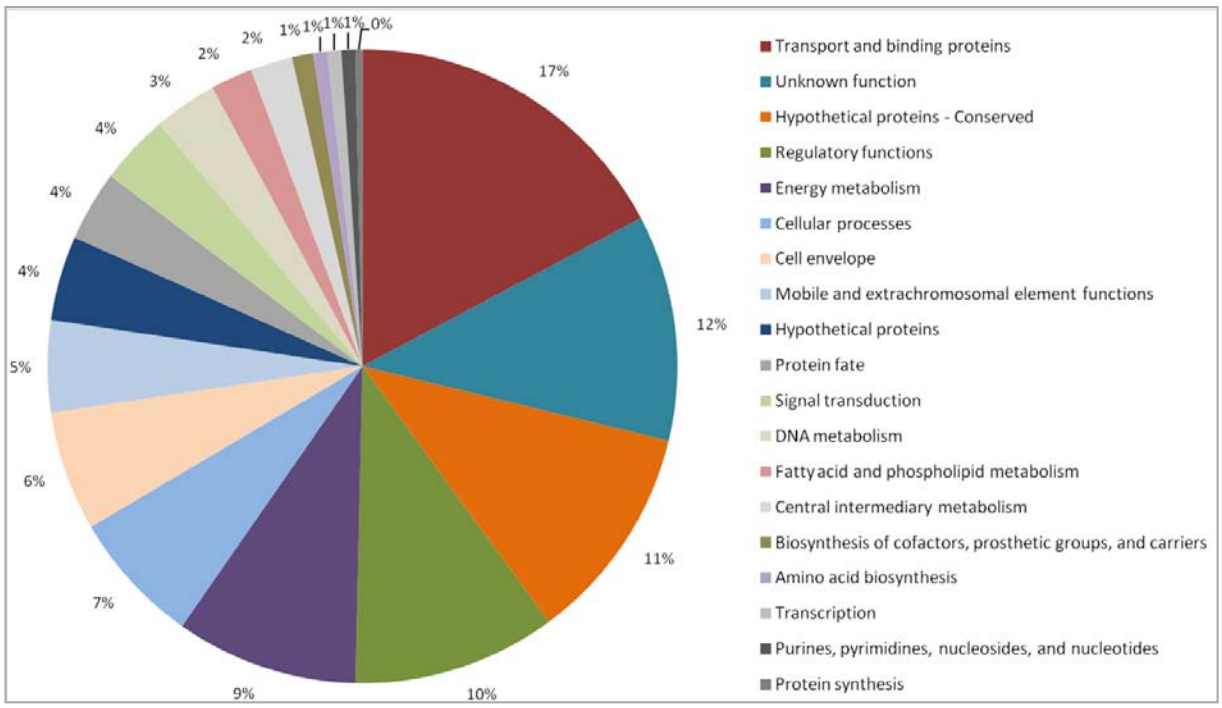



**Figure 4-8: Representation of the cellular role categories of the disrupted genes of *P. putida* TMT mutants**
The pie chart distributes the disrupted genes of *P. putida* TMT mutant into different cellular role categories. One colored sector represents one category. The percentage of genes composing this category is annotated outside the sector. The color of the sector corresponds to the one given in the legend on the right side. One cellular role category was not found in the mutant, thus not represented on the pie chart. The percentages are calculated out of 261 strains due to the presence of 28 encoded proteins which play a role in two different categories.

After separated transposition of both mini-Tn*5* derivatives and selection of single and double mutants, analysis of the disrupted genes indicated a major loss of encoded proteins playing roles in the three following categories: 'Transport and binding proteins', 'Regulatory functions' and 'Energy metabolism'. These categories corresponded to those most represented in the genome of *P. putida* KT2440 (http://cmr.jcvi.org/cgi-bin/CMR/shared/Menu.cgi?menu=genome; Comprehensive Microbial Resources by choosing the Role category Pie Chart for the desired organism). The percentage for each category was estimated out of the 5,437 genes annotated in the genome.

Among the genes targeted by a mini-Tn*5 TF* more than once, PP_0806 was extracted. It is interesting to note that the *lapF* gene was again the target of a mini-transposon (hit once by the mini-Tn*5 KpF*). In total after the establishment of both the libraries (*P. putida* SMT and TMT mutants), the gene was hit three times from independent transpositions.

Generation of single and double mutant of *P. putida* TEC1 allowed the generation of *P. putida* SMT and TMT strains. These strains provided information about genes which are not considered as being essential for the sustainability in M9 minimal medium supplemented with uracil and citrate as carbon source. Furthermore *P. putida* TMT mutants will be further used for the generation of a deleted mutant (*P. putida* $\Delta_x$) devoid of parts of its genome.

## 4.2 *Pseudomonas putida* $\Delta_x$ mutants: random deletion of genomic fragments in *Pseudomonas putida* TEC1

### 4.2.1 Generation of *Pseudomonas putida* $\Delta_1$ mutant strains

Due to the random insertion of both mini-Tn*5* derivatives into the genome of *P. putida* TEC1 different combinations were obtained for the mini-Tn*5 KpF*/mini-Tn*5 TF* pair of each *P. putida* TMT mutants. Based on the direction of these integrated mini-Tn*5* derivatives a pre-selection of potential candidates for deletion was established. Each *P. putida* TMT mutant from the library exhibited one mini-Tn*5 KpF* and one mini-Tn*5 TF* with defined directions. The "minus" and "positive" signs appearing in Table 8-6 highlighted three different cases; see Figure 4-9 below. However, only a single combination could lead to the deletion of a genomic fragment together with both resistance cassettes. This meant that a single *FRT* site, framed by the Iend and the Oend of the mini-Tn*5* derivatives, was left after the excision operated by the flippase and will participate to the following part of the experiments which will be described in section 4.2.2.1.1.

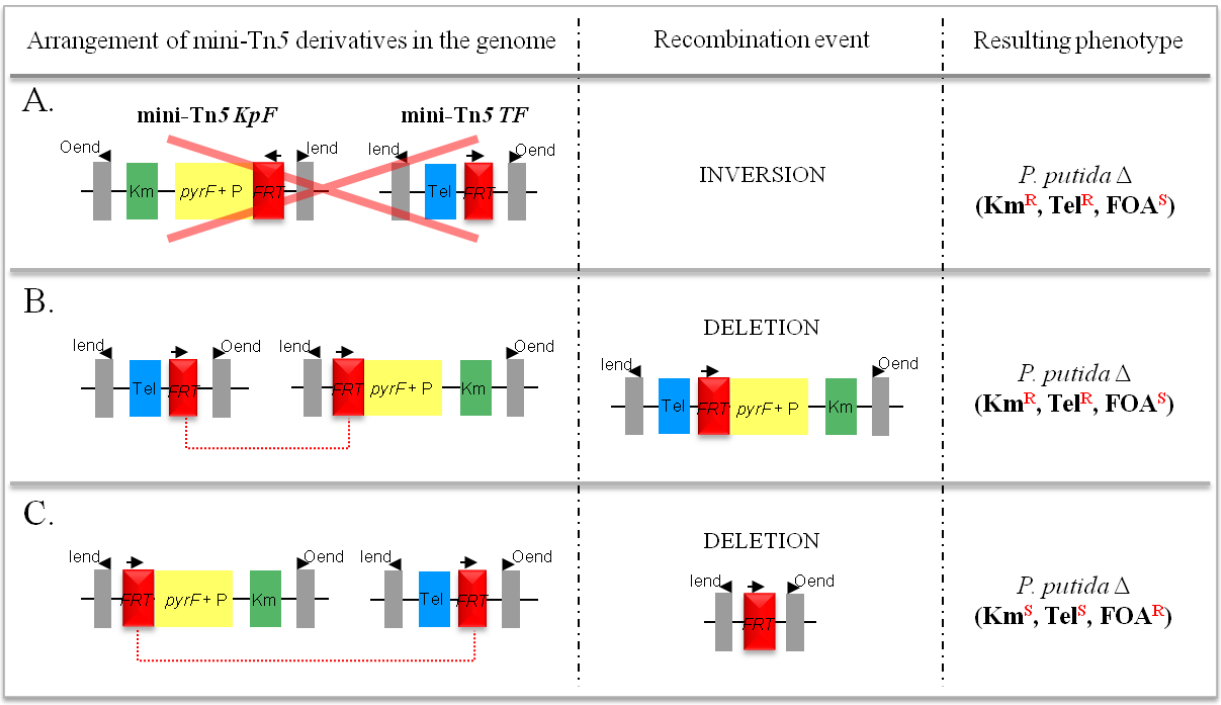



**Figure 4-9: Possible directions of *FRT* sites in a *P. putida* TMT strain and phenotype of the corresponding deleted mutant.**

**A.** The *FRT* sites are mapped in opposite direction. In this case the action of the Flipase would lead to an inversion of the genomic fragment situated between both mini-Tn*5* derivatives, leading to the same phenotype as the *P. putida* TMT strain. **B.** The *FRT* sites are mapped in the same direction provoking the deletion of the genomic fragment. However, due to the position of both mini-Tn*5 KpF* and mini-Tn*5 TF* in the chromosome the resistances remain in the genome after action of the Flipase (red dots line). The appearing phenotype is the same as for the *P. putida* TMT strain. **C.** Deletion of genomic fragment occurs along with the loss of the resistances after action of the Flipase (red dots line). This case is the only one leading to the desired phenotype. The red cross applied on case A indicates that if this configuration was found in the *P. putida* TMT library, mutants were not chosen as potential candidates for genomic deletion.

As previously explained in section 4.1.1.1, one *FRT* site should be oriented towards the Oend of mini-Tn*5 TF* and the other one towards the inner part of the mini-Tn*5 KpF*. This requirement did not affect the way of recombination (deletion or inversion of the genomic fragment) but rather caused the simultaneous loss of both resistances. As shown in Figure 4-9 when both mini-Tn*5* derivatives are positioned in opposite directions (A) only an inversion of the genomic fragment could occur, whereas in the two other cases, (B) and (C), a deletion event is physically possible. However, the 3$^{rd}$ recombination (C) was the only one to lead to the loss of the resistance traits.

#### 4.2.1.1 Inversion of genomic fragments after action of the Flp recombinase

When both mini-Tn*5* derivatives showed an opposite sign the corresponding *P. putida* TMT strains were not considered any longer as potential candidates for genomic deletion. As depicted for case A in Figure 4-9, when one mini-Tn*5* was indicated with a negative sign (*FRT* site in the same direction as the "-" strand of the genome) and the second mini-Tn*5* with a positive sign (*FRT* site in the same direction as the "+" strand), recombination between both *FRT* fragments should lead to the inversion of the genomic segment between the Flp recognition targets. In Table 8-6 the inversion cases are highlighted in gray. In total, 138 *P. putida* TMT mutants would lead to an inversion case, which represents 54.3 % of the library.

#### 4.2.1.2 Deletion of genomic fragments after action of the Flp recombinase

Deletion of genomic fragment in *P. putida* TEC1 required a defined position and orientation of both the mini-Tn*5* derivatives. The choice and the submission of 34 *P. putida* TMT mutants for the deletion step will be explained in the following paragraphs.

##### 4.2.1.2.1 Random determination of potential candidates for genomic deletion

The first attempts for the deletion of genomic fragments in *P. putida* TMT strains were initiated directly after the generation of the first double mutants. At this stage only a few *P. putida* TMT putative mutants were sequenced for the position of both mini-Tn*5* derivatives in their genome. Out of the 34 TMT mutants chosen for the establishment of the random deletion method, 16 revealed, unfortunately, the opposite orientation of both mini-transposons (case A) once the AP-PCR and the sequencing reactions were carried out. This latter information further explains why the first attempts on the 16 strains were not successful (data not shown).

Among the *P. putida* TMT mutants, TMT 12, TMT 21 and TMT 37 (Table 8-6) were selected for genomic deletion prior to the mapping of the mini-Tn*5 TF*. They were therefore randomly chosen for the deletion step before knowing that they belong to cases A and B in Figure 4-9. They will be mentioned in the section 4.2.1.2.2.

#### 4.2.1.2.2 Selection of potential candidates for genomic deletion

In order to confirm the random mutagenesis method associated with site-specific recombination, 116 mutants were found to belong to cases A and B (Figure 4-9) after mapping of the mini-Tn*5* derivatives. This value represented 45.7 % of all *P. putida* TMT mutants for which both mini-Tn*5* derivatives were mapped. *In silico* analysis revealed that from these 116 strains 53 should result in a {$Km^R$, $Tel^R$ and $FOA^S$} phenotype after the action of the Flp recombinase and 63 strains should lead to a {$Km^S$, $Tel^S$ and $FOA^R$} phenotype. These strains are highlighted in Table 8-6 by using a yellow and green, respectively. Only those mutants who could potentially lose both resistance traits were considered as the most interesting for further experiments.

To classify the potential candidates and establish an order of priority regarding the submission to the deletion step, these 63 candidates (24.8 %) representing case C were analyzed in greater detail. The size of the putative deleted fragments was estimated by subtracting the coordinates from the position of both the mini-Tn*5* derivatives. The predicted sizes ranged from ~41.5 kb (TMT 91) to ~2,985 kb (TMT 444) and are given for each *P. putida* TMT mutant belonging to cases A and B (see Table 8-6). In addition the position of both mini-Tn*5* derivatives in the genome was compared with the origin of replication (*ori*C) of the *P. putida* TEC1 strain. The deletion of the origin of replication was never done before. Therefore, the consequences of such a deletion would lead to different behavior, maybe to the death, of the cells. Thus, any candidate leading *in silico* to the removal of the *oriC* were not first tested, in the context of the establishment of the method. A last point was raised concerning the possible presence of essential genes in the region flanked by two *FRT* sites. Two predictions of gene essentiality were carried out with the *in silico* metabolic reconstruction of *P. putida* KT2440 in glucose minimal medium (Nogales *et al.*, 2008; Puchalka *et al.*, 2008). With an eye on these two models the fragments to be deleted were also considered depending on the essentiality prediction. Combining these three elements, a first target group of seven candidates (TMT 69, 91, 113, 253, 407, 475 and 477) was chosen (case C). Additionally, three more colonies (TMT 12, 21 and 37) were randomly selected within the random determination (section 4.2.1.2.1) and were further verified as belonging to case C. Owing these methods of selection, 10 candidates were used for the first attempts of deletion. The size of the genomic fragment to be deleted ranged between ~41.5 kb and ~1,023 kb (TMT 113) with two exceptions for TMT 12 and TMT 37, randomly chosen, which carry a fragment long of 2,457 kb and 1,959 kb, respectively. The main aim of this study was the establishment of the method itself, therefore only those candidates which had the highest chance of success for the site-specific recombination and genomic deletion were considered. For this purpose and in order to increase the chances of obtaining the first deleted mutant, eight additional candidates were selected from case B using the same criterion: the size of the fragments to be deleted was framed by the size of TMT 53 (~46 kb) and TMT 550 (~613 kb). It should be noted that none of the candidates chosen from cases B and C were predicted to lose the origin of replication.

#### 4.2.1.2.3 Deletion experiments on the selected potential *P. putida* TMT candidates

The selected double mutants were submitted separately to the last steps of the deletion procedure, namely the insertion of pBBFLP vector and the expression of the Flp recombinase leading to the loss of the genomic fragment. As described in section 3.3.5.5 the pBBFLP vector was transferred to the mutant strains by triparental mating. The presence of the vector was selected on M9 minimal medium supplemented with uracil, citrate and tetracycline. Parallel inoculation overnight of the $Tc^R$ exconjugants in LB broth and M9 minimal medium both supplemented with cit, ura and Tc allowed the recombination to occur between the two *FRT* sites present in the genome. In order to remove the pBBFLP plasmid from the cells after recombination the inocula were plated on both media containing sucrose (5 % w/v). For each deletion case up to several hundred of single colonies (from both the LB and M9 plate) were picked and streaked onto four different LB and M9 plates. The first of these four plates was supplemented with citrate and uracil in order to collect biomass and the other three contained kanamycin, tellurite or FOA for sensitivity/resistance screening. Finally, if no successful mutant was obtained, the screening was conducted with the next *P. putida* TMT candidate.

The TMT candidates from case B were not expected to lose any resistance after deletion. Therefore, several colonies were streaked on different media supplemented with Km and Tel, in order to confirm their presence in the chromosome. The occurrence of the deletion was screened by PCR amplification using primer sets targeting the mini-Tn*5* derivative (kilA rev and Oend-Km For2, see Tables 8-2 and 8-4). Unfortunately, the expected fragment of 1.1 kb was not detected for any of the candidates from case B data not shown, indicating that the deletion did not occur.

Only two of the candidates from case C showed the desired phenotype: $Km^S$, $Tel^S$ and $FOA^R$. These deleted mutants were named *P. putida* $\Delta_1$-91 (originally from *P. putida* TMT 91) and *P. putida* $\Delta_1$-407 (originally from *P. putida* TMT 407). For *P. putida* $\Delta_1$-91, colonies were obtained only on LB plates supplemented with sucrose. Interestingly, for *P. putida* $\Delta_1$-407 colonies were obtained on LB but also on parallel plates prepared with M9 minimal medium, even though they appeared extremely slowly. Ten individual colonies were selected for each deleted mutant from the LB + ura + cit Agar plate and inoculated into LB broth supplemented with uracil and citrate in order to collect and conserve the knockout mutants.

### 4.2.1.3 Confirmation of the deletions

The deletions for both knockout mutants were confirmed by three successive steps: phenotypic characterization, Southern blot hybridization and AP-PCR/sequencing experiments. The results will be shown separately for *P. putida* $\Delta_1$-91 (section 4.2.1.3.1) and *P. putida* $\Delta_1$-407 (section 4.2.1.3.2) mutants.

#### 4.2.1.3.1 *Pseudomonas putida* $\Delta_1$-91 mutant strain

The sensitivity to both antibiotics was the first indication of genomic rearrangement after action of the Flp recombination enzyme. Resistance to FOA, however, confirmed the presence of the $Km^S$ phenotype. The Km cassette and *pyrF* operon were always kept together, if one was deleted the second one was automatically removed. The plates comprising the mutants are discussed below and can be observed in Figure 4-10.

- **Phenotypic characterization**

After triparental mating of *P. putida* TMT 91 with plasmid pBBFLP and subsequent plating on sucrose medium, colonies were screened on the four different selection plates for the desired phenotype, see Figure 4-10.

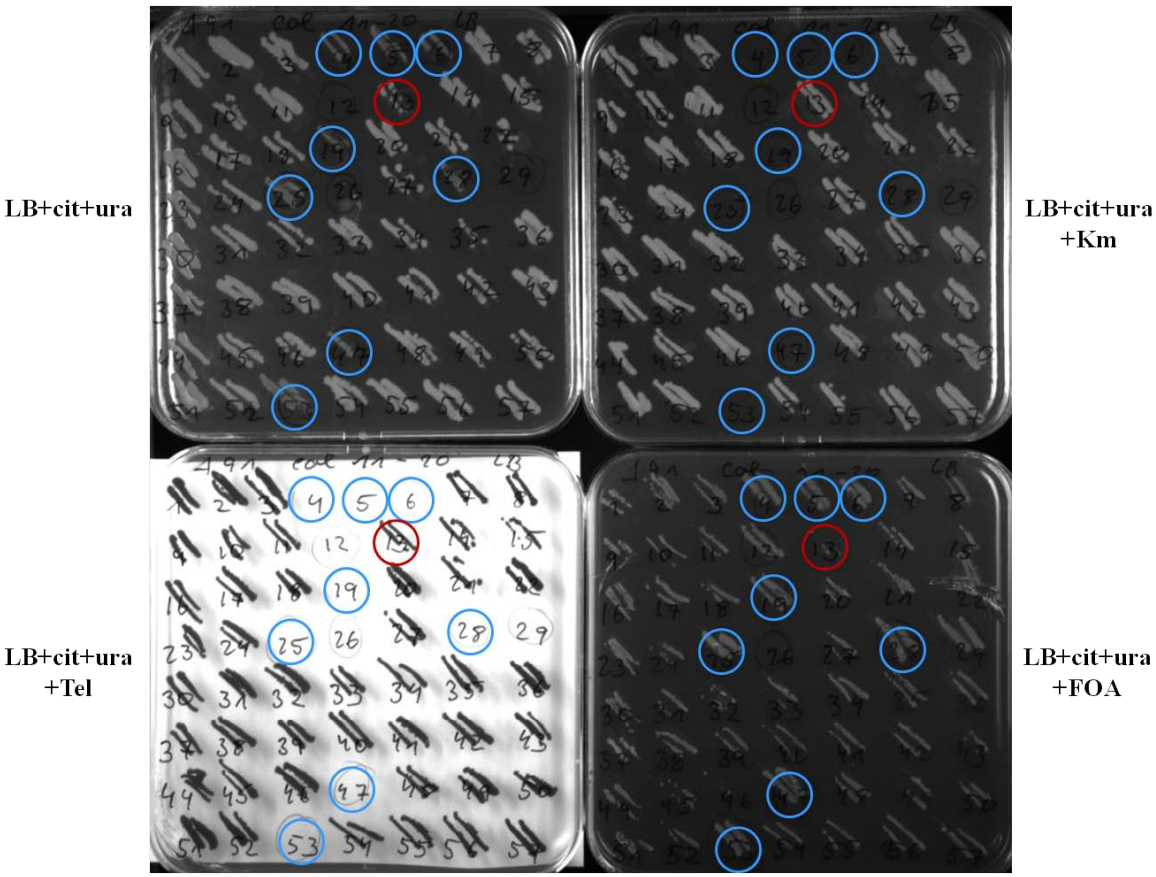



**Figure 4-10: Phenotypic selection of putative *P. putida* $\Delta_1$-91 mutants after action of the Flp recombinase.** The composition of the media is described on the side of each square plate. Eight colonies circled in blue (*P. putida* $\Delta_1$-91 4, 5, 6, 19, 25, 28, 47 and 53) show a $Km^S$, $Tel^S$ and $FOA^R$ phenotype, indicating a genomic rearrangement. The col-13, circled in red, was chosen as negative control ($Km^R$, $Tel^R$ and $FOA^S$ phenotype).

Colony 13, red circle in Figure 4-10, displayed the same phenotype as the *P. putida* TMT 91 strain ($Km^R$, $Tel^R$ and $FOA^S$) indicating the absence of recombination between the two *FRT* sites inserted in the genome. Accordingly, colony 13 was selected as control for the following experiments. The eight

following colonies, *P. putida* $\Delta_1$-91 col. 4, 5, 6, 19, 25, 28, 47 and 53, blue circles, were selected for further verifications due to their $Km^S$, $Tel^S$ and $FOA^R$ phenotype, corresponding to case C (Figure 4-9).

- **Southern blot hybridization**

*P. putida* $\Delta_1$-91 col. 4, 5, 19 and 25 were chosen out of the eight putative mutants previously highlighted for the Southern blot experiments. The gDNA from these colonies as well as gDNA from *P. putida* TEC1 (negative control) and from *P. putida* TMT 91 and *P. putida* $\Delta_1$-91 col-13 (positive controls) was digested and loaded on an agarose gel, see Figure 4-11.

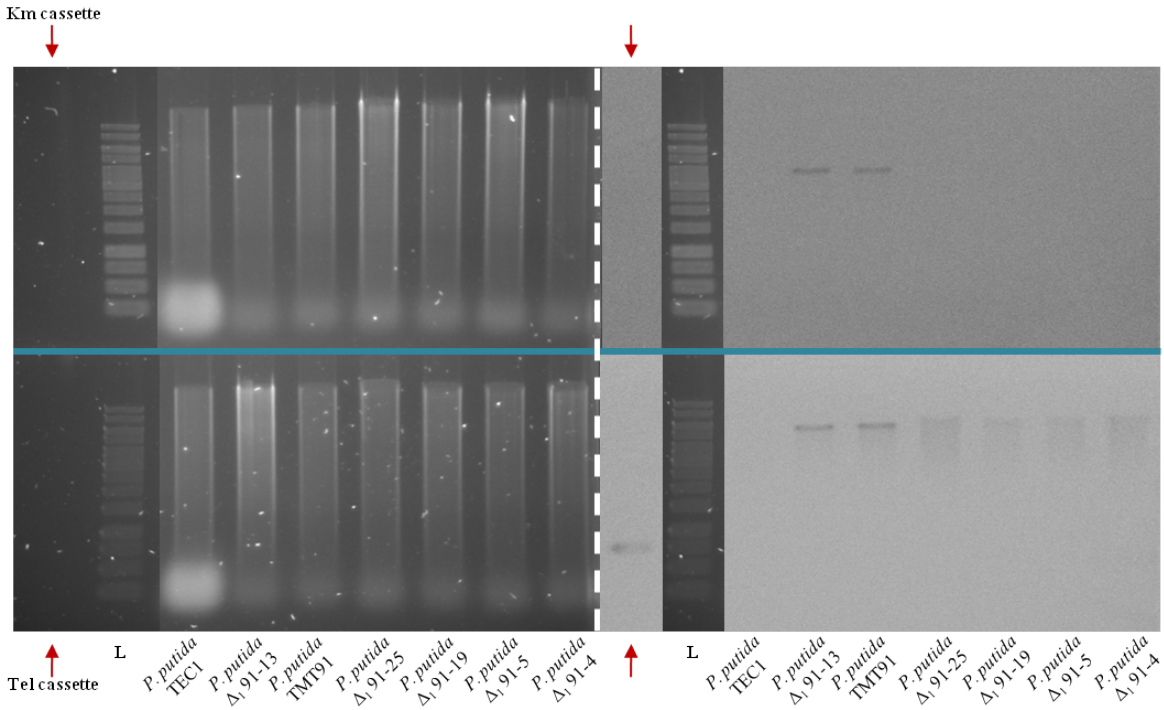



**Figure 4-11: Verification of the putative *P. putida* $\Delta_1$-91 mutants by Southern blot experiments.**
Agarose gels (left side of the dashed white line) and corresponding blots (right side of the dashed white line) of the *Eco*RV restricted gDNA from *P. putida* TEC1 (negative control), *P. putida* TMT-91 and *P. putida* $\Delta_1$-91 col.13, as two positive controls, and *P. putida* $\Delta_1$-91col. 4, 5, 19 and 25, as indicated at the bottom of the Figure. Above the blue line: hybridization with the labeled Km probe. Below the blue line: hybridization with the labeled Tel probe. The ladder (L) loaded on the gel is manually reported on the picture of the blot to simplify the reading of the size of the bands.

As expected, two bands (of ~5.5 and 4 kb in size) were detected for the two positive controls *P. putida* TMT 91 and the *P. putida* $\Delta_1$-91 col. 13, indicating the insertion of the mini-Tn*5 TF* into the genome and the presence of the mini-Tn*5 KpF*, respectively. No signal was detected for the negative control (TEC1) (Figure 4-11). This confirms that no recombination event occurred for col. 13. After hybridization with the Km probe no other signal was found for any of the putative mutant strains, col. 4 to 25, which validated the phenotypic results. On the Tel blot faint bands which were slightly

larger in size than the positive controls were detected for four of the colonies (col. 4, 5, 19 and 25). Due to the presence of these signals further verification was therefore required to confirm the validity of the four selected putative deleted mutants.

- **PCR amplification**

The scheme of the deletion step in *P. putida* TMT 91 is shown in Figure 4-12. To further confirm the deletion after action of the flippase, the data from the library based on the AP-PCR and sequencing results were used to design primers in order to amplify from outside the Iend of mini-Tn*5 KpF* and the Oend of mini-Tn*5 TF*. PP_3490 and PP_3529 were the two corresponding genes disrupted by the mini-Tn*5* derivatives (Figure 4-12). Primers PP_3490 fw and PP_3529 rv (Table 8-4) were used to amplify the newly formed segment which should have appeared after deletion and recombination in *P. putida* $\Delta_1$-91 mutants.

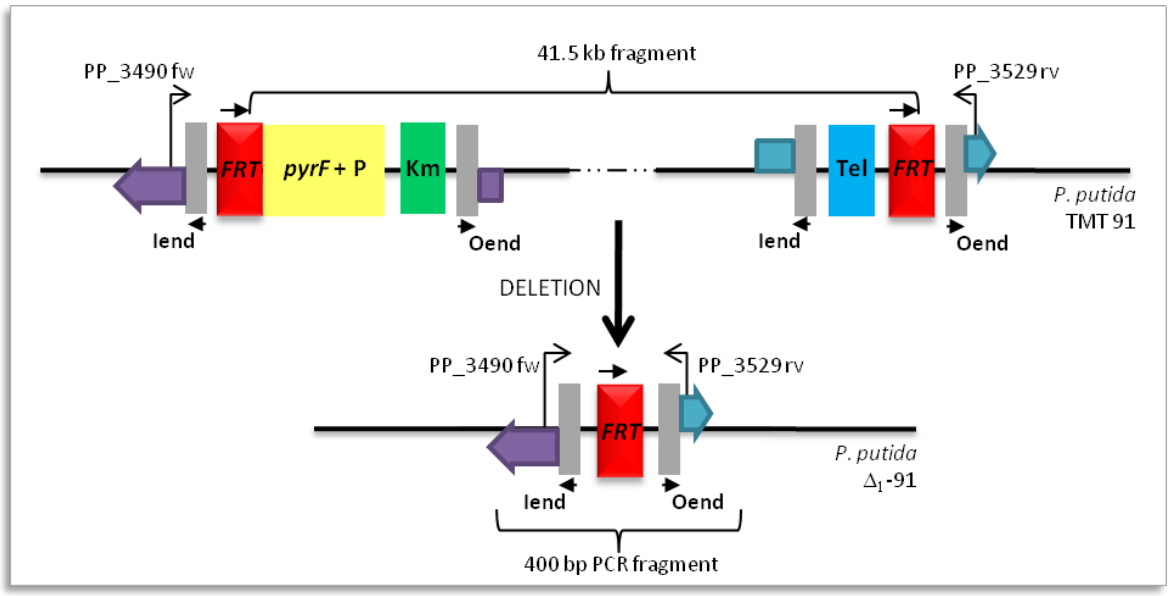



**Figure 4-12: Schematic representation of the random deletion of a genomic fragment in *P. putida* $\Delta_1$-91.**
At the top: representation of the chromosome of *P. putida* TMT 91 strain carrying both mini-Tn5 derivatives. The lilac arrow represents the disrupted PP_3490 gene. The blue arrow represents the disrupted PP_3529 gene. The predicted size of the fragment before deletion is reported above the brace. Position of the primers used for the PCR reaction is indicated outside the minitransposons. At the bottom: settings after deletion of the genomic fragment. One Iend (from mini-Tn*5 KpF*), one Oend (from mini-Tn*5 TF*) and one *FRT* site (recombined from the previous target sites) composing the so-called "new borne cassette" of the *P. putida* $\Delta_1$-91, framed by one part of the PP_3490 gene (lilac arrow) and one part of the PP_3529 gene (blue arrow) are left after recombination.

PCR amplification was performed using the gDNA of four putative deleted mutants (col. 4, 5, 19 and 25) and three different controls (*P. putida* TEC1, TMT 91 and $\Delta_1$-91 col. 13). A band of approximately 400 bp was predicted to appear only if recombination between both *FRT* sites occurred. Although some non-specific annealing of the primers was also observed in the reactions, a

strong band of the expected size of ~400 bp was detected for each of the four putative mutants (Figure 4-12). Amplicon bands were excised, purified and sequenced as described in section 3.3.3.

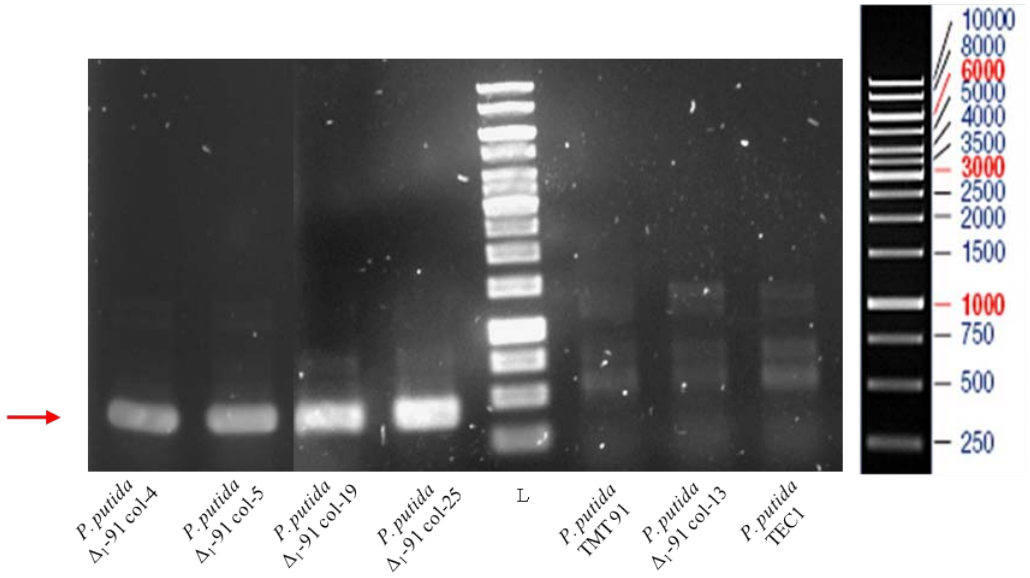



**Figure 4-13: PCR amplification of the new DNA junction created after deletion in *P. putida* $\Delta_1$-91 mutants.**
Four putative deleted mutants were submitted to PCR with PP_3490 fw and PP_3529 rv primers. The same PCR was carried out with *P. putida* TEC1, *P. putida* TMT 91 and *P. putida* $\Delta_1$-91 col. 13 as negative controls. The lane L represents the 1 kb ladder. The red arrow indicates the expected fragment.

- **Sequencing experiments and final confirmation**

Each purified PCR product was sequenced from both ends. After alignment with the genomic sequence of *P. putida* KT2440 and the different mini-Tn*5* carrying plasmids the results confirmed the genomic deletion for all the putative mutants. The sequencing results for *P. putida* $\Delta_1$-91 col. 5, taken as an example, highlight a number of different points:

- Insertions of both mini-Tn*5* were correctly predicted *in silico* at the base pair level with the AP-PCR/sequencing method for the establishment of the library. Indeed the sequence obtained for *P. putida* $\Delta_1$-91 col. 5, situated upstream of the Iend of mini-Tn*5* *KpF* and downstream of the Oend of mini-Tn*5* *TF* corresponded exactly to the predicted coordinates found.
- Mutations were found neither in the flanking segments nor in the {Iend-*FRT*-Oend} sequence of the newly formed cassette.

The scheme of the newly formed cassette formed after action of the Flp recombinase is depicted in Figure 4-12. The sequencing result is detailed in Figure 4-14. The alignment was performed with the

genes of interest from *P. putida* KT2440 (PP_3490 and PP_3529) and both mini-Tn*5* *KpF* and *TF*. The sequence obtained was started from the PP_3490 side.

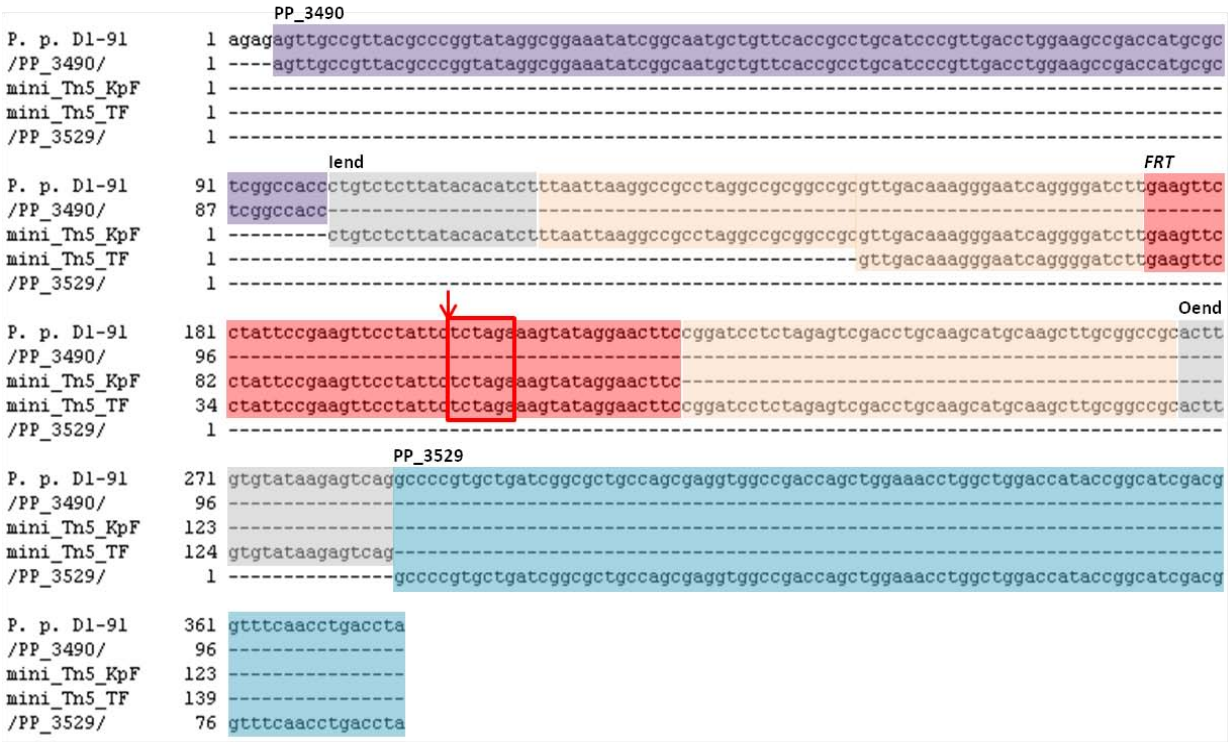



**Figure 4-14: Nucleotide alignment of the new DNA junction with different DNA fragments for the *P. putida* $\Delta_1$-91 mutant.**
Four sequences were aligned with the new DNA junction created after deletion and recombination in the *P. putida* $\Delta_1$-91 mutant. The sequence of the new junction corresponds to the one obtained after PCR with PP_3490 fw and PP_3529 rv. The new junction is named as P. p. D1-91 on the scheme and the disrupted framing genes are symbolized by the name of each gene inscribed between slashes on the left side of the figure. The mini-Tn*5* derivatives' sequences are indicated by the name of each mini-Tn*5*. After alignment, matching between two or more sequences is symbolized by colored boxes. **In lilac**: part of PP_3490 upstream the new junction; **in gray**: the two ends (Iend and Oend); **in red**: *FRT* site with the *Xba*I site highlighted by a red rectangle and the site of recombination marked with a red arrow; **in blue**: part of PP_3529 downstream the new junction; **in orange**: fragments corresponding to the sequence of the mini-Tn*5* derivatives situated between the ends and the *FRT* site. The appropriated legends are reported above each rectangle.

Among all the tested *P. putida* $\Delta_1$-91 putative mutants, a single colony was selected to participate in further experiments. *P. putida* $\Delta_1$-91 col. 5 was confirmed as the first mutant with a single genomic deletion created by random mutagenesis and site-directed homologous recombination. The deletion generated a 0.67 % reduction of the chromosome from the original *P. putida* TEC1 strain. Colony 5 was subsequently selected for further experiments as detailed further.

#### 4.2.1.3.2 *Pseudomonas putida* $\Delta_1$-407 mutant strain

- **Phenotypic characterization**

After insertion of pBBFLP via triparental mating only two single colonies were obtained on M9 agar supplemented with ura, cit and Tc. The exconjugants were inoculated exclusively in LB medium, which was supplemented either with ura, cit and Tc or with ura, cit, Tc and FOA. Addition of FOA prevented the growth of the bacteria, however, they were able to grow (albeit slowly) in its absence, reaching an appropriate cell density after five days. After plating 50 µl of the culture on sucrose supplemented medium, 44 colonies were detected and all of them were used for the verification of the phenotype, see Figure 4-15. Four different LB agar plates were subsequently used for streaking of the colonies supplemented either with Km, Tel or FOA or without any supplement.

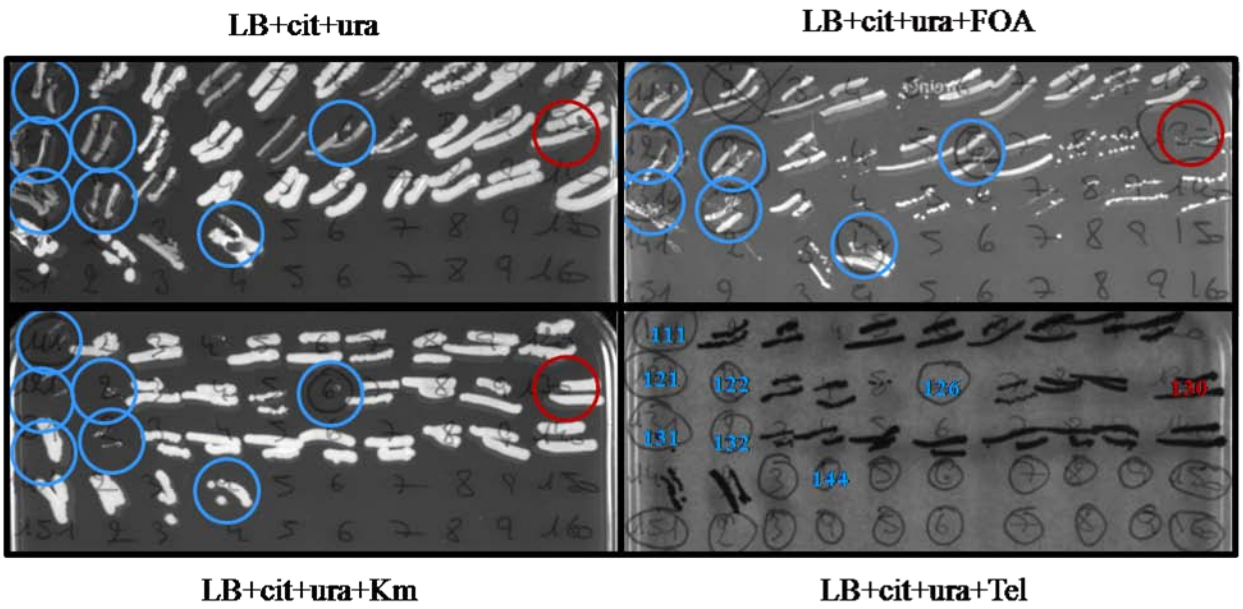



**Figure 4-15: Phenotypic selection of putative *P. putida* $\Delta_1$-407 mutants after action of the Flp recombinase.**
Seven colonies circled in blue color (111, 121, 122, 126, 131, 132 and 144) show two different interesting phenotypes: $Km^S$, $Tel^S$ and $FOA^R$ and $Km^R$, $Tel^S$ and $FOA^R$, indicating a genomic rearrangement. The col. 130, circled in red, indicated the one chosen as negative control ($Km^R$, $Tel^R$ and $FOA^S$ phenotype).

Five putative *P. putida* $\Delta_1$-407 colonies (col. 111, 121, 122, 126 and 132) showed the following expected phenotype: $Km^S$, $Tel^S$ and $FOA^R$, corresponding to case C, Figure 4-9. They were, thus, selected for further analyses. Col. 131 and 144 showed a particular phenotype like $Km^R$, $Tel^S$ and $FOA^R$, see Figure 4-15, which does not correspond to any of the possible combination cases (Figure 4-9). Whilst sensitivity to tellurite was obvious; the colonies were able to grow on LB agar supplemented with Km and on FOA. In order to substantiate this result, PCR (described in section 3.3.3) was used to amplify the <Km::*pyrF*::*FRT*> cassette from the gDNA of the two colonies. The primers Km-pBAM F and Iend Rev2 (see Tables 8-3 and 8-4) were used in the reaction. Controls

were carried out with the gDNA from *P. putida* TEC1, *P. putida* TMT 407 and the five positive colonies selected (col. 111, 121, 122, 126 and 132). The result of the PCR is shown in Figure 4-16.

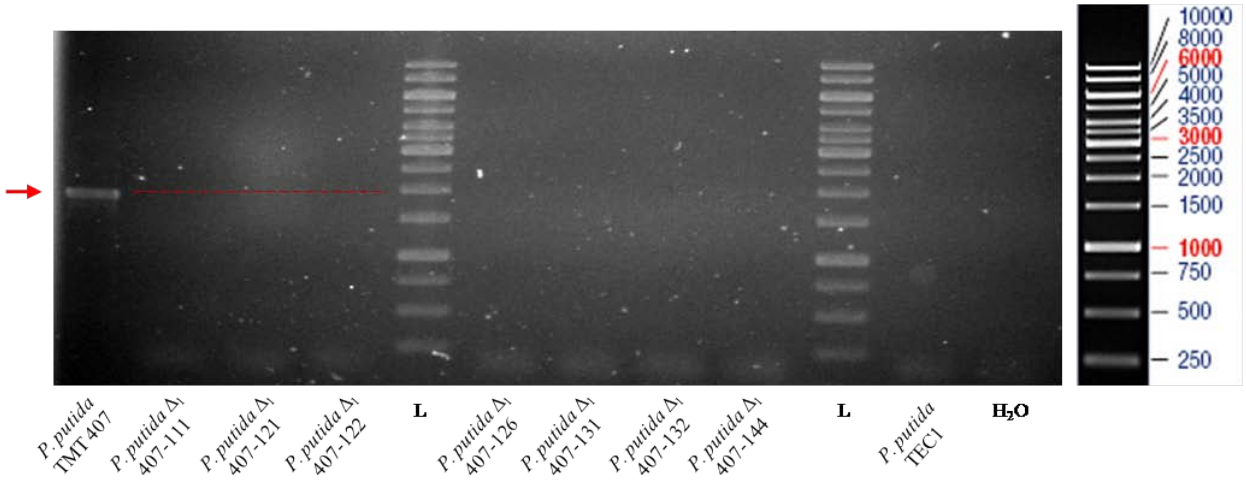



**Figure 4-16: Verification of the putative *P. putida* $\Delta_1$-407 mutants by PCR amplification of the *KpF* cassette.**
The amplification of the <Km::*pyrF*::*FRT*> fragment for the different *P. putida* $\Delta_1$-407 colonies obtained after deletion and the three controls: $H_2O$, *P. putida* TEC1 and *P. putida* TMT 407, was loaded on a gel. The ladder (L) was loaded twice on the gel for a better visualization of the band sizes. The red arrow indicates the sole band appearing on the gel corresponding to the expected size (~1,900 bp) for the fragment in the genome of *P. putida* TMT 407, positive control.

As can be seen in Figure 4-16, the *P. putida* $\Delta_1$-407 colonies 111, 121, 122, 126 and 132 were confirmed by PCR for the loss of the mini-Tn*5 KpF* as amplicon could not be detected in comparison to the positive control. This was similarly the case for col. 131 and 144. No PCR fragment for <Km::*pyrF*::*FRT*> was revealed on the gel, indicating that the loss of the mini-Tn*5 KpF* should have occurred as well for these two colonies. Further verifications were therefore carried out with all seven colonies along with col. 130, which was used as a negative control.

- **Southern blot hybridization**

The genomic DNA of the seven putative *P. putida* $\Delta_1$-407 mutants (col. 111, 121, 122, 126, 131, 132 and 144) as well as the gDNA of the two controls *P. putida* TEC1 and *P. putida* $\Delta_1$-407 col. 130 were used for the blotting experiments. The amplification of the probes was based on the two following plasmids: pBAM1 and pUT/*TF*, employed as positive control. The results are illustrated in Figure 4-17.

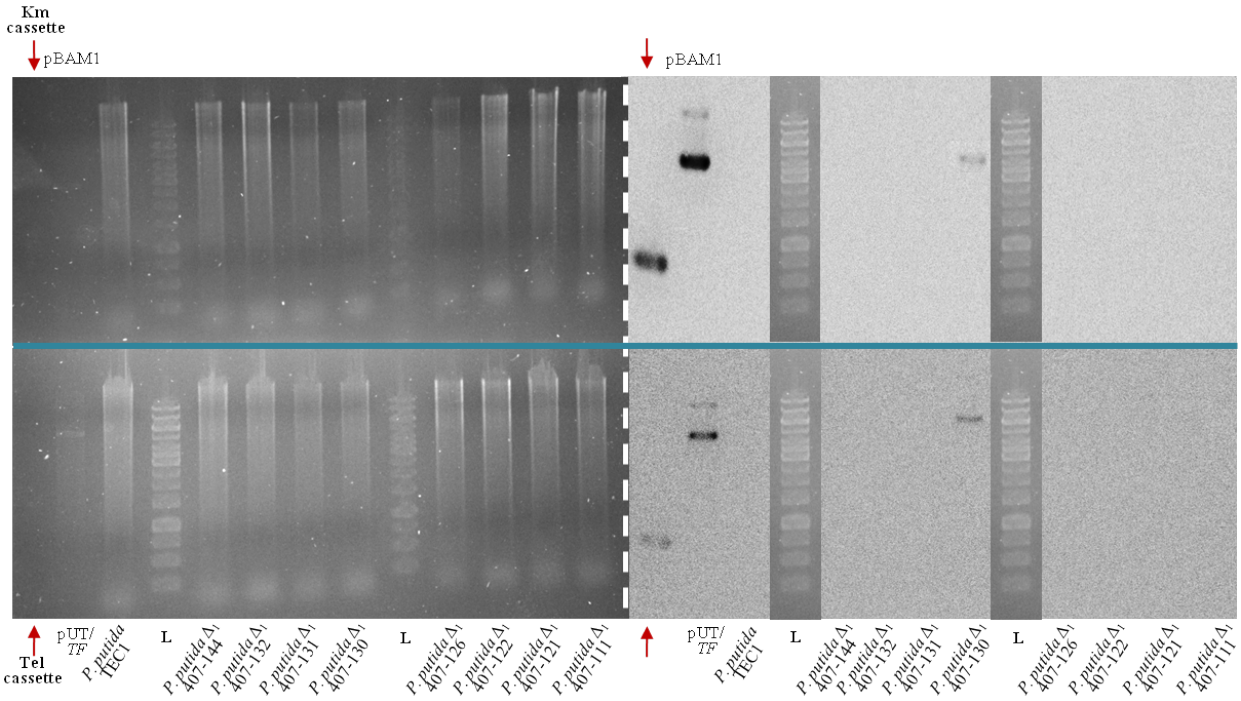



**Figure 4-17: Verification of the putative *P. putida* $\Delta_1$-407 mutants by Southern blot experiments.**
Agarose gels (left side of the dashed white line) and corresponding blots (right side of the dashed white line) of the *Eco*RV restricted gDNA from *P. putida* TEC1 (negative control), *P. putida* $\Delta_1$-407 col-130 (positive control) and *P. putida* $\Delta_1$-407 col. 111, 121, 122, 126, 131, 132 and 144. Above the blue line: hybridization with labeled Km probe. The pBAM1 vector is used as a positive control. The red arrow indicates the Km cassette. Below the blue line: hybridization with labeled Tel probe. The pUT/*TF* vector is used as positive control. The red arrow indicates the Tel cassette. The ladders (L) present on the agarose gel are reported manually on the blot picture for a more convenient read of the bands size.

Although the expected sizes of the plasmids are 4,394 bp (pBAM1) and 4,201 bp (pUT/TF) two bands were seen for each vector after hybridization indicating the different conformation of the DNA. Both PCR fragments of the expected size of 760 bp (Km) and 749 bp (Tel) were detected. No hybridization was observed with the gDNA from the negative control *P. putida* TEC1 wild-type strain. The single band found on the picture for both blots corresponded to the hybridization of the Km and the Tel probes with the gDNA of *P. putida* $\Delta_1$-407 col.130. The presence of both mini-Tn*5* derivatives in its genome indicated that no deletion occurred for this colony. No signal was found for any of the putative *P. putida* $\Delta_1$-407 strains. The seven colonies were confirmed for not carrying any mini-Tn*5* derivative in their genome after hybridization.

- **AP-PCR amplification and sequencing experiments**

In the case of *P. putida* $\Delta_1$-407, the two genes PP_3534 and PP_3733 were disrupted by the mini-Tn*5* derivatives. In order to amplify the deleted region, primers PP_3534 fw and PP_3733 rv were designed upstream of the Iend of mini-Tn*5 KpF* and downstream of the Oend of mini-Tn*5 TF*. The scheme of the deletion and recombination occurring in *P. putida* TMT 407 is depicted in Figure 4-18.

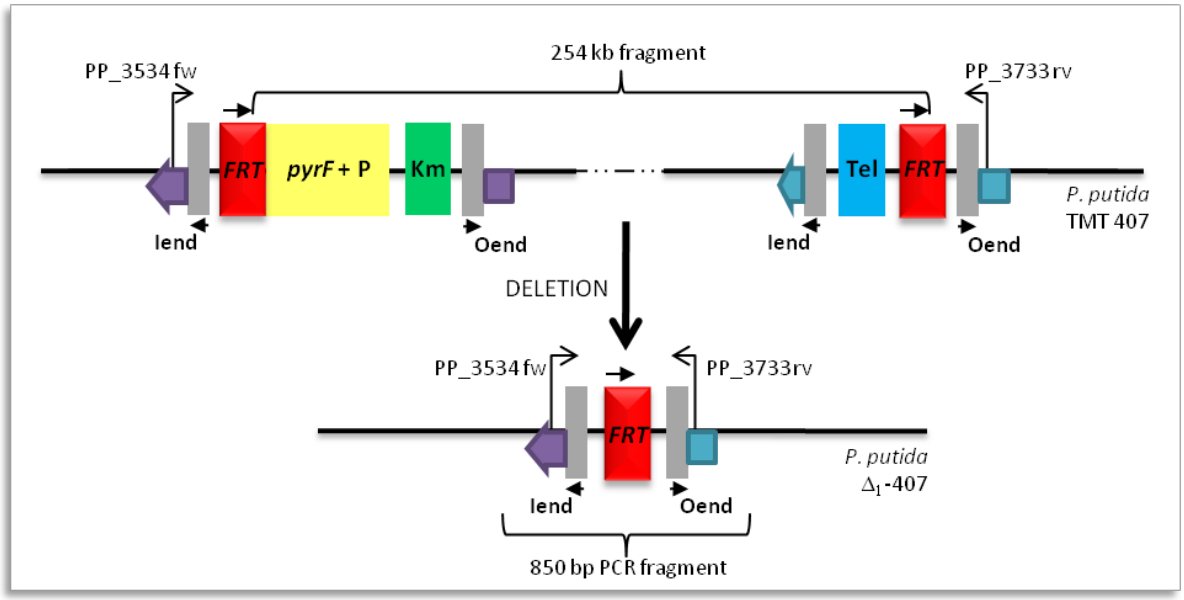



**Figure 4-18: Schematic representation of the random deletion of a genomic fragment in *P. putida* $\Delta_1$-407.** At the top: representation of the chromosome of *P. putida* TMT 407 strain carrying both mini-Tn5 derivatives. The lilac arrow represents the disrupted PP_3534 gene. The blue arrow represents the disrupted PP_3733 gene. The predicted size of the fragment after deletion is reported above the largest brace. The position of the primers used for the PCR reaction is indicated outside the minitransposons. At the bottom: after deletion of the genomic fragment one Iend (from mini-Tn*5 KpF*), one Oend (from mini-Tn*5 TF*) and one *FRT* site (recombined from the previous target sites), composing the so-called "new borne cassette" of the *P. putida* $\Delta_1$-407, framed by one part of the PP_3534 gene (lilac arrow) and one part of the PP_3733 gene (blue arrow) are left after recombination.

The gDNA from seven putative deleted mutants as well as from *P. putida* mutants $\Delta_1$-407 col-130, TMT 407 and strain TEC1 was used as template for subsequent PCR reactions. Four microliters of each PCR product were loaded on the gel shown in Figure 4-19. The expected size of the fragment containing a part of the disrupted genes and the {Iend-*FRT*-Oend} fragment was expected to be ~850 bp, as shown in Figure 4-18.

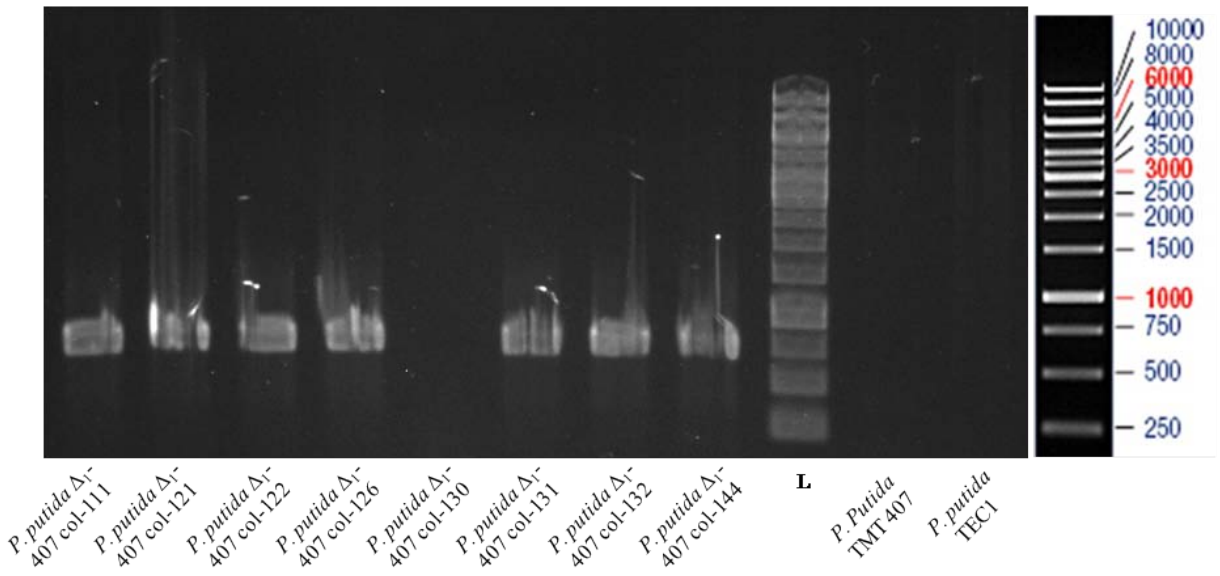



**Figure 4-19: PCR amplification of the new DNA junction created after deletion in *P. putida* $\Delta_1$-407 mutants.**
Seven putative deleted mutants were submitted to PCR with PP_3534 fw and PP_3733 rv primers. The same PCR was done for *P. putida* TEC1, *P. putida* TMT 407 and *P. putida* $\Delta_1$-407 col. 130 as negative controls. The lane L represents the 1 kb ladder.

A single band with the expected size was revealed on the gel for each of the seven putative mutants whereas no band was found for the negative controls. The PCR products were purified directly after the reaction and sequenced from both ends. After alignment with the genome of *P. putida* KT2440 and with the mini-Tn*5* carrying plasmids, the putative mutants were confirmed as deleted mutants. The result for *P. putida* $\Delta_1$-407 col-121 is presented in Figure 4-20, after sequencing with the PP_3534 fw primer. The analysis of the alignment highlights a number of different points:

- Insertion of the mini-Tn*5* derivatives was predicted with accuracy during the AP-PCR and sequencing experiments
- No single point mutation was found in the whole fragment containing the new DNA junction between PP_3534 and PP_3733

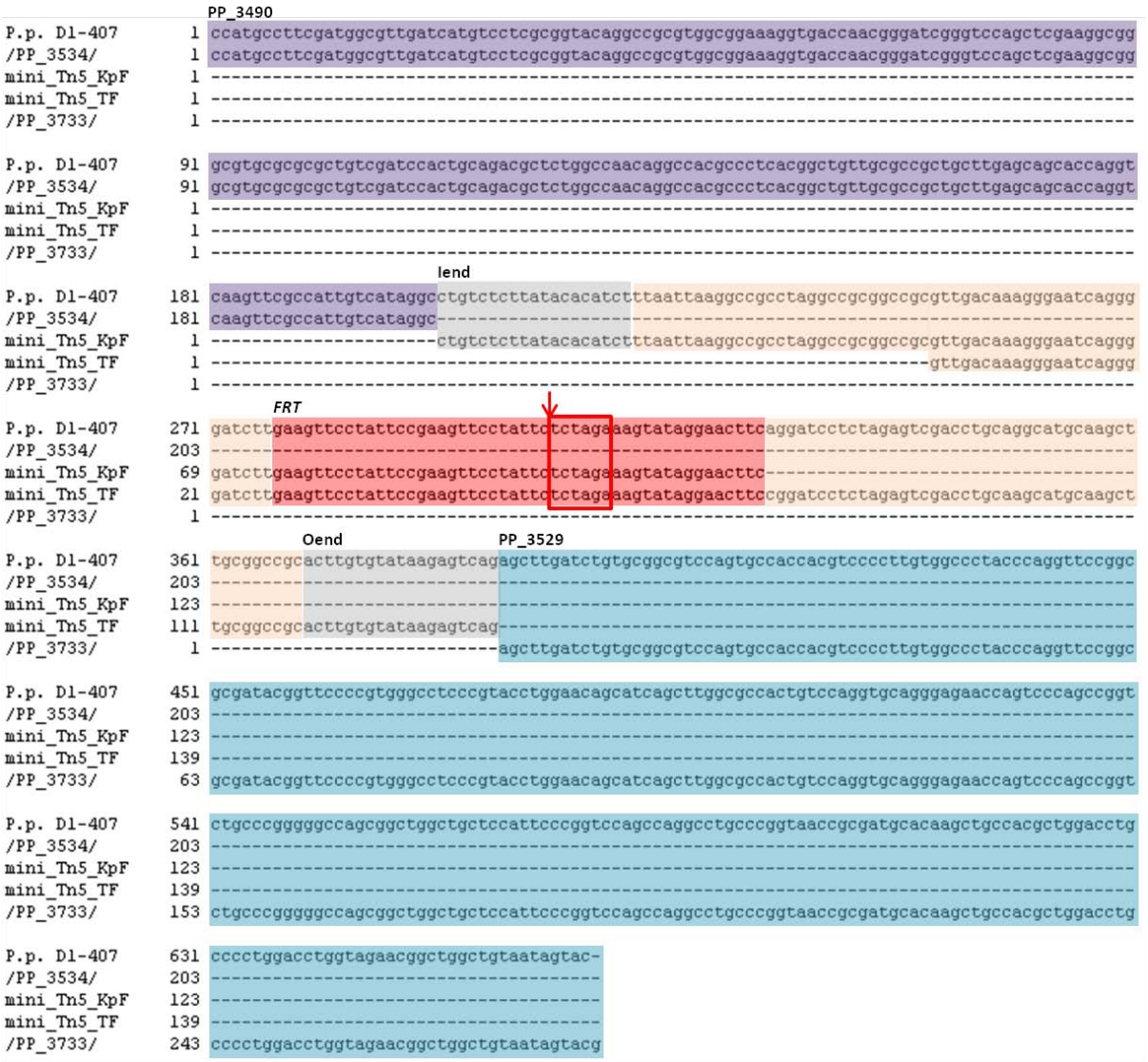



**Figure 4-20: Nucleotide alignment of the new DNA junction with different DNA fragments for the *P. putida* $\Delta_1$-407 mutant.**

Four sequences were aligned with the new DNA junction created after deletion and recombination in the *P. putida* $\Delta_1$-91 mutant. The sequence of the new junction corresponds to the one obtained after PCR with primers PP_3534 fw and PP_3733 rv. The new junction is named as P. p. D1-407 on the scheme and the disrupted framing genes are symbolized by the name of each gene between slashes, on the left side of the figure. The mini-Tn*5* derivatives' sequences are indicated by the name of each mini-Tn*5*. After alignment, matching between two or more sequences is symbolized by colored boxes. **In lilac**: part of PP_3534 upstream the new junction; **in gray**: the two ends (Iend and Oend); **in red**: *FRT* site with the *Xba*I site highlighted by a red rectangle and the site of recombination marked with a red arrow; **in blue**: part of PP_3733 downstream the new junction; **in orange**: fragments corresponding to the sequence of the mini-Tn*5* derivatives situated between the ends and the *FRT* site. The appropriated legends are reported above each rectangle.

The sequencing results confirmed with 100 % of similarity the loss of the genomic fragment. Thus *P. putida* $\Delta_1$-407 col. 121 was chosen to represent the second mutant strain with a single genomic deletion obtained after random mutagenesis and site-directed homologous recombination. For this strain the percentage of reduction in a single cell reached 4.1 % of the size of the original *P. putida* TEC1 genome; a reduction greater than observed for the first *P. putida* $\Delta_1$-91 mutant.

In summary, Southern blot experiments confirmed the presence of both mini-Tn*5* derivatives in the genomes of *P. putida* $\Delta_1$-91 col. 13 and *P. putida* $\Delta_1$-407 col. 130. Even though the deletion did not succeed in these two colonies both were kept under the designation of "$\Delta_1$" in the following results but always in association with "positive / negative control", when compared with the putative deleted mutants. All the controls were approved by the different blots created for the Km and Tel cassettes.

#### 4.2.1.4 Characterization of the two first *Pseudomonas putida* $\Delta_1$ mutant strains

Sequencing of the fragment carrying the recombinant genomic DNA of both *P. putida* $\Delta_1$-91 and $\Delta_1$-407 mutants contributed to the first characterization of the strains. In order to get more information about the function of the genes from the deleted region, sequence data were evaluated across different databases (section 3.4.4.2). The product names were principally found with KEGG. Whilst much of the data was similar regardless of the database used, a few genes differed when compared between KEGG and CMR. In this situation the orthologous genes in three other *P. putida* strains (namely *P. putida* F1, GB-1 and W619) were analyzed with KEGG and CMR in order to find the greatest similarity. The cellular role categories were defined by the CMR. In instances where the gene encoding a protein was related to two or more categories, it was considered in both cases as detailed further below. For each mutant the whole list of deleted genes is described in the following way: loci IDs are given as well as the corresponding gene, product name, KEGG Orthology (KO) identifier (including the K number) and/or Enzyme Commission number (EC) and cellular role. For a better reading of the table, the rows corresponding to proteins encoded by genes playing the same cellular role category were highlighted with the same color. When a protein happened to play several roles in the cell, the rows were left uncolored.

##### 4.2.1.4.1 *Pseudomonas putida* $\Delta_1$-91 mutant strain

As previously mentioned, the size of the chromosome from the mutant strain was reduced by 0.6 %, corresponding to the ~41.5 kb predicted *in silico*. Both flanking genes PP_3490 and PP_3529 were disrupted by the mini-Tn*5 KpF* and mini-Tn*5 TF*, respectively, and part of them was fully removed after action of the flippase (as shown in Figure 4-12). In addition to these two partially knockouts 38 other genes were entirely removed from the genome and are listed in Table 8-8.

- **Gene list and cellular role categories**

The 40 deleted genes were classified by their involvement in different cellular role categories. The different groups are represented in Figure 4-21. In total eight genes (PP_3490, PP_3494-PP_3495, PP_3504-PP_3505, PP_3518, PP_3520 and PP_3525) encode conserved hypothetical proteins which, as the largest category (see Table 8-8 and Figure 4-21), represents 20 % of all knock-out ORFs. Six genes (PP_3493, PP_3503, PP_3513, PP_3516, PP_3526 and PP_3527) encode different

transcriptional regulators and belonged, therefore, to the category ‘Regulatory function’ (DNA or protein interactions). This was the most represented of the cellular role category with a total of 15 % ORFs in the mutant. Five genes (PP_3498 to 3502) were found to encode different transposases (6,1 % of the total amount of transposases) and thus belong to the ‘Mobile and extrachromosomal element functions’ category. Equal representations (7.5 %) of the three following cellular role categories were found for ‘Biosynthesis of cofactors, prosthetic groups and carriers’, ‘Amino acid biosynthesis’ and ‘Cell envelope’. Three protein coding genes composed each group: (PP_3506, PP_3507, PP_3508), (PP_3511, PP_3414, PP_3515) and (PP_3512, PP_3519, PP_3521), respectively. Proteins related to the ‘Transport and binding of proteins’, ‘Fatty acid and phospholipids metabolism’ and ‘Transcription’ were also equally represented (5 %) with two genes for each category (PP_3491, PP_3492), (PP_3506, PP_3528) and (PP_3496, PP_3522), respectively. The last defined cellular role category accounted for 2.5 % of the ORFs (see Table 8-8) and consisted of one gene (PP_3497) coding for a U32 family peptidase related to ‘Protein fate’. The rest of the genes were shared between four hypothetical protein-coding genes (PP_3510, PP_3517, PP_3523 and PP_3524), which were not associated to any cellular role category (10 %) and two genes coding for proteins with unknown functions (PP_3509 and PP_3529) (5 %). In summary, nine defined cellular role categories were involved by the genomic deletion out of the 12 found in *P. putida* $\Delta_1$-91. Thirty-five percent (17 genes) of the whole encoded proteins were not fully described yet. However, none of these genes appeared to be essential for any metabolic or cellular process under the described conditions.

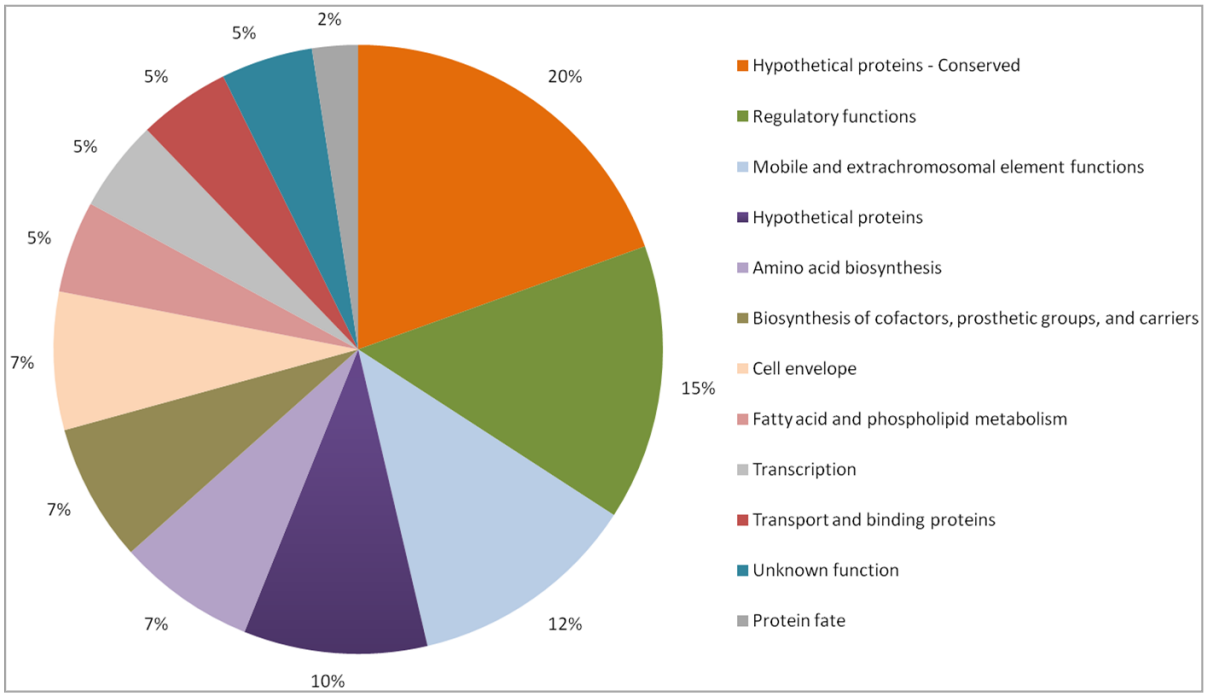



**Figure 4-21: Representation of the cellular categories concerned by the genomic deletion in *P. putida* $\Delta_1$-91.**

The pie chart distributes the deleted genes of *P. putida* $\Delta_1$-91 mutant into different cellular role categories. One colored sector represents one category. The percentage of genes composing this category is given outside the sector. The color of the sector corresponds to the one given in the legend on the right side. Eight cellular role categories (out of the 20 present in *P. putida* KT2440) were not found in the mutant, thus not represented on the pie chart. The percentages are calculated out of 41 strains due to the presence of one encoded protein which plays a role in two different categories.

Among the 40 knockout genes, five code for transposases situated downstream of each other. The 'Mobile and extrachromosomal element function' category was composed of five of these protein coding genes. Three transposases formed the operon ISPpu14. Another operon, predicted from the Database of prokaryotic Operons (DOOR, (Mao *et al.*, 2009) and (Dam *et al.*, 2007)), was found to be composed of the genes PP_3506, PP_3507 (*cobN*) and PP_3508 (*cobW*). *CobN* and *cobW* belong to the genes encoding enzymes participating in the biosynthesis of vitamin B12. They are present in single copy within the genome. However it was previously described by Molina-Henares and colleagues that mutants carrying a mini-transposon in one of those genes did not require vitamin B12 for growth in M9 minimal medium. They presumed that from precorrin-2, which is formed in an intermediate step leading to the synthesis of vitamin B12, another pathway must exist for vitamin B12 biosynthesis (Molina-Henares *et al.*, 2010). The *P. putida* $\Delta_1$-91 mutant was therefore not expected to be a vitamin B12 auxotroph.

- **Verification of gene essentiality**

One focus of the genomic deletion is the removal of PP_3511 (*ilvE*), a gene that encodes the branched-chain amino acid aminotransferase, present in a single copy in the genome. This gene was

predicted to be essential for growth in M9 minimal medium supplemented with glucose as a sole carbon source (Nogales *et al.*, 2008). During the screening of the genome-wide mutant library, obtained by transposition with mini-Tn*5* and growth in M9 minimal medium with citrate as the sole carbon source, no IlvE mutant was found with an auxotrophy for valine or leucine (Molina-Henares *et al.*, 2010). *P. putida* $\Delta_1$-91 was found here to be unable to grow on M9 minimal medium supplemented with citrate (data not shown). A hypothesis was formulated regarding the essentiality of the *ilvE* gene for the growth of the mutant in the defined conditions. Therefore, growth experiments were carried out comparing the behavior of the mutant with the wild-type strain in M9 minimal medium supplemented with uracil, citrate and different combinations of amino acids in the case of the mutant. In previous studies it was shown that growth of *P. putida* on L-valine was stimulated by traces of L-Isoleucine, thus added in the medium (0.005 %) for the growth experiments (Marshall and Sokatch, 1972). The general aspect of the growth curves are shown in Figure 4-22. The experiments were run over 23 h and repeated twice. Each mutant was cultivated in duplicate.

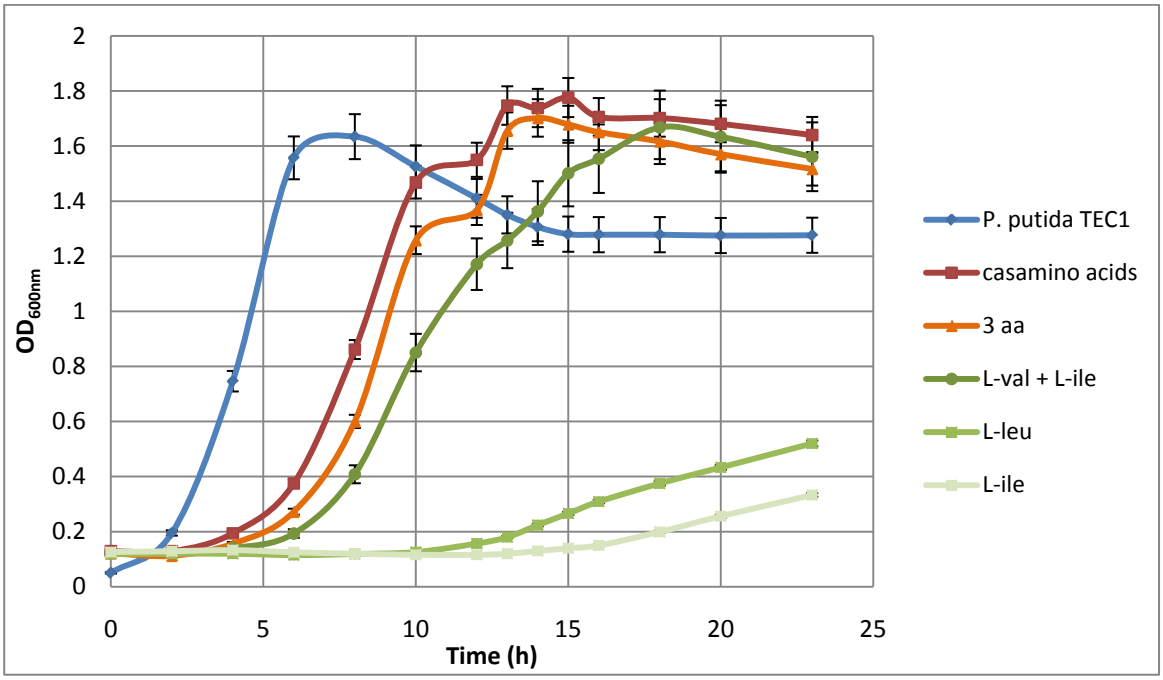



**Figure 4-22: Growth comparison of *P. putida* $\Delta_1$-91 mutant and *P. putida* TEC1 in M9 minimal medium.** Measurement of the optical density ($OD_{600nm}$) against the time (h) shows the behavior of *P. putida* $\Delta_1$-91 under different conditions in comparison with the wild-type strain (in M9 medium only with uracil and citrate). **In dark blue**: *P. putida* TEC1 without amino acid supplement. **In red**: M9 minimal medium is supplemented with citrate as carbon source, uracil and casamino acids (Casaa). **In dark green**: the casamino acids are replaced by the three branched-chain amino acids. **In green**: M9 supplemented with citrate, uracil and L-isoleucine only. **In violet**: L-isoleucine is replaced by L-leucine. **In light green**: the main amino acid supplemented is L-valine; some traces of L-isoleucine are present too.

Addition of the casamino acids (all amino acids but tryptophan) or only the three branched-chain amino acids was sufficient for the mutant to recover its initial growth to a level almost comparable to that of the wild-type, although the lag phase was considerably longer (Figure 4-22). Interestingly, in

presence of L-valine and traces of L-isoleucine the cells show the same tendency of growth as with more amino acids present in the medium. Separate supplementation of L-leucine or L-isoleucine increases the lag phase of the growth and does not seem to sustain the cells as efficiently as with L-valine (+ 0.005 % L-isoleucine).

The presence of amino acid(s) in the medium allows the strain to recover its growth. In order to confirm the *ilvE* gene as single essential gene provoking the non-growth of the bacteria in minimal medium only supplemented with uracil and citrate, complementation test was carried out. For this purpose pBBR1MCS-5 vector carrying the wild-type PP_3511 gene was used as described in paragraph 3.3.7. The modified plasmid was transferred by electroporation to *P. putida* $\Delta_1$-91 and replication of the vector was selected on M9 minimal medium supplemented with citrate, uracil and gentamycin and without any source of amino acid. After overnight incubation colonies were grown on the plates meaning that the deleted mutant recovered its ability to grow on minimal medium with citrate as carbon source.

#### 4.2.1.4.2 *Pseudomonas putida* $\Delta_1$-407 mutant strain

Generating the second mutant *P. putida* $\Delta_1$-407, a more substantial reduction in the genome was obtained. A total of 198 genes were removed which, corresponding to 253,851 bp, reduced the genome by 4.1 % of its total size. In addition, the two "framing" ORFs, PP_3534 and PP_3733, were disrupted due to the insertion of the mini-Tn*5 KpF* and mini-Tn*5 TF*, respectively. These 200 genes are listed in Table 8-9. The data were extracted from the same databases as mentioned previously, paragraph 4.2.1.4.

- **Description of the gene list**

The genomic deletion in *P. putida* $\Delta_1$-407 mutant was much greater than that which occurred in *P. putida* $\Delta_1$-91 and therewith a larger number of cellular role categories were involved. In total, 15 roles were defined (Table 8-9) of which the two predominant categories were 'Transport and binding proteins' (38 genes, 18 % of the total number deleted) and 'Regulatory functions' (35 genes, 16 % of the total number deleted), see Figure 4-23. Each of the following 12 categories represented less than 10 % of all functions. 'Energy metabolism' comprised 20 genes (9 %) coding for proteins playing such a role. The following nine categories were represented by nine genes or less (≤ 5 %). 'Cellular processes', 'Fatty acid and phospholipid metabolism', 'Signal transduction', 'Amino acid biosynthesis' and 'Mobile and extrachromosomal element functions' comprise nine to four encoded proteins (4 to 2 %). 'Cell envelope', 'Biosynthesis of cofactors, prosthetic groups, and carriers', 'Protein fate', 'Transcription', 'Central intermediary metabolism' and 'DNA metabolism' were slightly affected by the deletion with only three genes for the first category, two for the following three and only one gene for the last two categories (less than 1 % and therefore noted as 0 % in Figure

4-23). Furthermore, two genes coding for proteins playing a role in 'Mobile and extrachromosomal element functions' were found with a disrupted reading frame.

Besides the defined categories, 42 proteins were annotated as 'Hypothetical proteins', forming the major group of categories (19 %). Nineteen genes were associated with a name but the encoded protein belonged to the 'Unknown function' category. Finally 15 genes encoded proteins conserved as hypothetical. In summary, 35 % of the cellular role categories affected by the knockout genes in *P. putida* $\Delta_1$-407 were not fully described yet, this is exactly the same percentage found for the genomic deletion in *P. putida* $\Delta_1$-91.

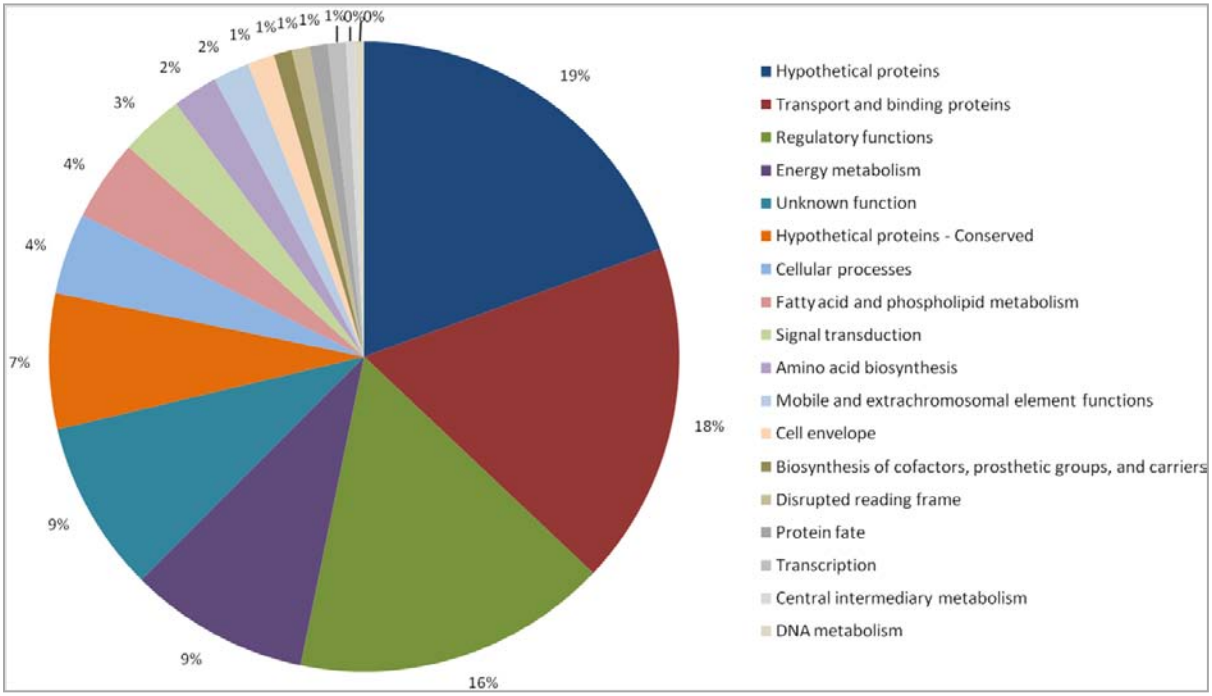



**Figure 4-23: Representation of the cellular role categories concerned by the genomic deletion in the *P. putida* $\Delta_1$-407 mutant.**

The pie chart distributes the gene knockouts of *P. putida* $\Delta_1$-407 mutant into different cellular role categories. One colored sector represents one category. The percentage of genes composing this category is annotated outside the sector. The color of the sector corresponds to the one given in the legend on the right. Two out of 20 cellular role categories were not found in the mutant, thus not represented on the pie chart. The percentages are calculated out of 216 strains due to the presence of 16 encoded proteins which play a role in two different categories.

The generation of knockout mutants devoid of one part of their genome proved that the random mutagenesis/site-specific recombination system is effective for *P. putida* TEC1 strain. The assessment of its reliability is further made with the generation of a second deletion in the same strain providing a good candidate for future biotechnological applications.

### 4.2.2 Generation of *Pseudomonas putida* $\Delta_2$ mutant strains

Random deletion of a genomic fragment in *P. putida* TEC1 was validated with the creation of the two first independent *P. putida* $\Delta_1$ mutants. However in order to follow the goal of the PhD work and to demonstrate the reliability of the deletion method, a second round of deletion was applied to the two *P. putida* $\Delta_1$. The following sections describe the creation of double-deletion mutants (strains where two consecutive genomic deletions occurred) along with the improvement of the deletion technique.

#### 4.2.2.1 Second genomic deletion in a single strain

Genomic deletion in *P. putida* TEC1 generated two deleted mutants, *P. putida* $\Delta_1$-91 and *P. putida* $\Delta_1$-407, in which a particular part of the genome was excised. The Flp recombinase generated a new DNA junction after deletion, which was sealed by a scar composed of two ends of the mini-transposons and a single *FRT* site with a defined orientation. This feature of the newly formed junction, in each mutant, is important for the following part of the experiments, namely the new insertion of mini-Tn*5* derivatives in these reduced genomes.

##### 4.2.2.1.1 Second application of the established random mutagenesis method

As being the first mutant obtained, *P. putida* $\Delta_1$-91 was used to test the second random deletion in an already reduced genome. For this purpose the same method was applied as described in section 4.2.1, meaning that the random mutagenesis by inserting mini-transposons and site-specific recombination was repeated. New disrupted mutants, named *P. putida* $\Delta_1$-91 SMT, were generated. The exconjugants were collected only in LB media due to the inability to grow of the $\Delta_1$-91 mutant in M9 medium. Each conjugation step evolved the presence of Nalidixic acid in the selective medium in order to counterselect the different *E. coli* cells (Goss *et al.*, 1965). Due to the presence of one FRT site in the genome, deletion could occur after insertion of each mini-Tn*5* derivative. The description of the different attempts for the deletion is given in sections 4.2.2.1.1.1 and 4.2.2.1.1.2.

###### *4.2.2.1.1.1 Deletion attempt after mini-Tn5 TF insertion*

The mini-Tn*5 TF* was the first inserted in the *P. putida* $\Delta_1$-91 mutant. In total, 45 *P. putida* $\Delta_1$-91 SMT colonies were tested and selected for their $Tel^R$ and $Pip^S$ phenotype, subjected to Southern blot hybridization (verification of a single mini-Tn*5* insertion) and used for AP-PCR and sequencing analysis in order to map the position of the newly inserted mini-transposon. Due to the presence of a single *FRT* site in the genome of $\Delta_1$-91, the mini-Tn*5 TF* insertion already provided potential candidates for genomic deletion. The configuration of this situation is shown in the second row of Table 4-1. A list was generated for mutations in *P. putida* $\Delta_1$-91 SMT (data not shown). The different mutants were analyzed in order to find potential candidates for deletion. By comparison of the orientation of the newly inserted mini-Tn*5 TF* with the orientation of the *FRT* site left in the genome

of *P. putida* $\Delta_1$-91, the $\Delta_1$-91 SMT mutants were categorized into three cases described in Figure 4-9. *In silico* estimation of the size of the fragment to be deleted facilitated the choice of potential candidates.

Three mutants were selected from case B (resistance to Tel remains) and three others from case C (loss of the Tel resistance) (Figure 4-9). The predicted sizes varied from 125 to 2,516 kb for case B and 664 to 1,077 kb for case C. The colonies were pooled together and subjected to the insertion of the pBBFLP vector into the cells and to following steps to obtain recombinant mutants (section 3.5). For the colonies from case C, verification of the deletion was done on LB plates containing uracil, citrate and tellurite. Unfortunately, none of the mutants tested lost the tellurite resistance. Since the resistance was not predicted to be lost after deletion in colonies from case C, verification of the genomic excision was carried out by amplification with specific primers following the same PCR as in section 4.2.1.2.3. However, none of the reactions generated a signal on the gel (data not shown), meaning that the fragment to amplify was too large and thus the deletion did not occur.

#### *4.2.2.1.1.2 Deletion attempt after insertion of the second mini-Tn5 derivative*

Ten *P. putida* $\Delta_1$-91 SMT mutants were pooled and subjected to the insertion of the second minit-Tn*5* derivative and generated *P. putida* $\Delta_1$-91 TMT mutants. The exconjugants were tested and confirmed for their $Tel^R$, $Km^R$, $FOA^S$ and $Pip^S$ phenotype. AP-PCR followed by sequencing was applied to the *P. putida* $\Delta_1$-91 TMT to map both mini-Tn*5* derivatives present in the genome. The different possibilities of organization of the three *FRT* sites are summarized in Table 4-1. Only in seven cases (8, 10, 13, 17, 20, 26 and 28) deletion of the genomic fragment and simultaneous loss of both resistance cassettes are possible. The *P. putida* $\Delta_1$-91 TMT mutants were screened *in silico* for one of these combinations in order to initiate the final step of deletion. Mutants were pre-selected and pooled to serve as acceptor strains in the conjugation involving *E. coli* HB101 and *E. coli* CC118λ*pir* pBBFLP. Deleted mutants were screened on different LB media supplemented with uracil and citrate and separately containing tellurite, kanamycin or fluoroorotic acid. Unfortunately the colonies appeared with the $Km^R$ and $Tel^R$ phenotype, indicating that no recombinant mutants were found.

**Table 4-1: Different combinations involving the three *FRT* sites present in the genome of *P. putida* $\Delta_1$ TMT strains.**

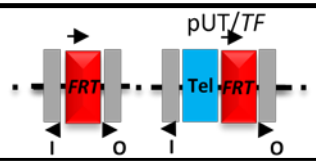
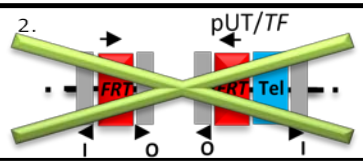
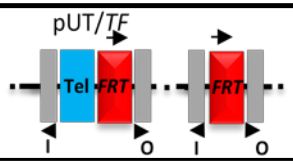
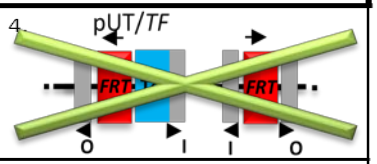
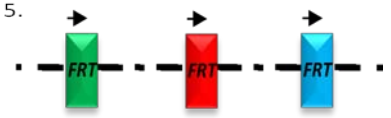
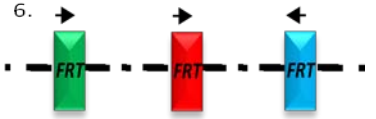
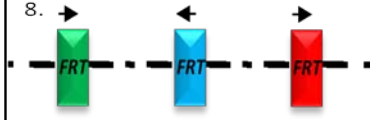
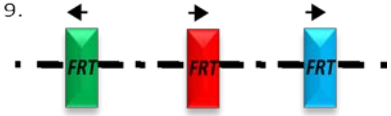
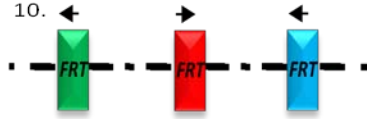
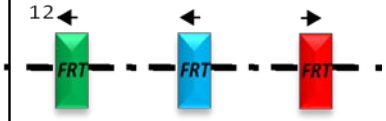
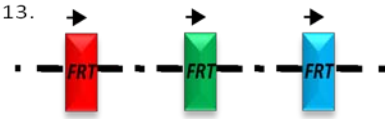
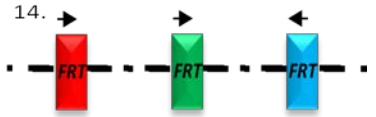
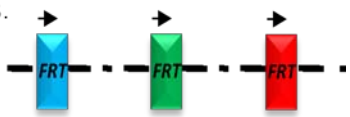
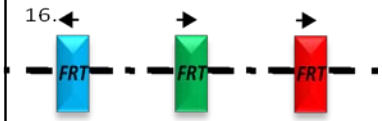
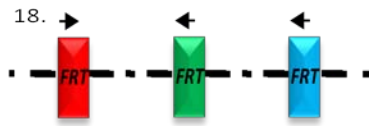
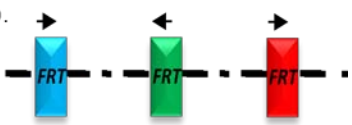
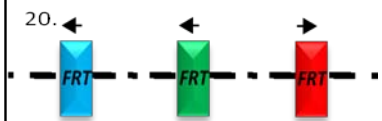
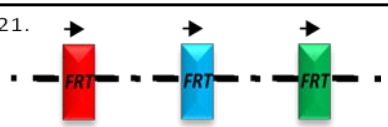
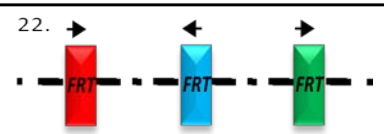
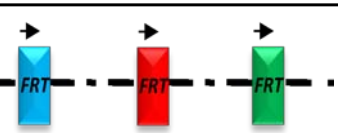
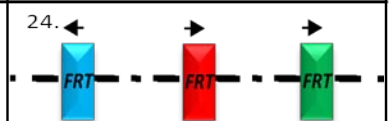
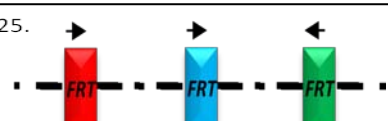
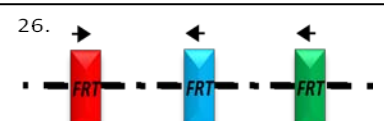
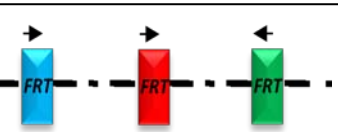
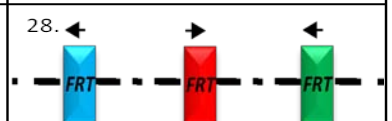
| mini-Tn*5 TF* / mini-Tn*5 KpF* | Sign | **mini-Tn*5 TF*** downstream of the scar (+) | **mini-Tn*5 TF*** downstream of the scar (-) | **mini-Tn*5 TF*** upstream of the scar (+) | **mini-Tn*5 TF*** upstream of the scar (-) |
|---|---|---|---|---|---|
| no **mini-Tn*5 KpF*** inserted | | 1.   | 2.   | 3.   | 4.   |
| **mini-Tn*5 KpF*** upstream of both the scar and the mini-Tn*5 TF* FRT | (+) | 5.   | 6.   | 7.   | 8.   |
| | (-) | 9.   | 10.   | 11.   | 12.   |
| **mini-Tn*5 KpF*** between the scar and the mini-Tn*5 TF* FRT | (+) | 13.   | 14.   | 15.   | 16.   |
| | (-) | 17.   | 18.   | 19.   | 20.   |
| **mini-Tn*5 KpF*** downstream of both the scar and the mini-Tn*5 TF* FRT | (+) | 21.   | 22.   | 23.   | 24.   |
| | (-) | 25.   | 26.   | 27.   | 28.   |

**Table 4-1: Different combinations involving the three *FRT* sites present in the genome of *P. putida* $\Delta_1$ TMT strains.**

All possible positions for both mini-Tn*5* derivatives are presented in the table. The four main columns indicate the orientation of mini-Tn*5 TF* (positive or negative sign) and its position in the genome in comparison with the new DNA junction (Iend-*FRT*-Oend, in red) of *P. putida* $\Delta_1$-91. The different rows indicate the position of mini-Tn*5 KpF* in the genome, either both fragments being downstream (new DNA junction and mini-Tn*5 TF*), in the middle or both being located upstream. The orientation of the "red" *FRT* site remains the same through the whole Table. The second row of the table (cases 1 to 4) summarizes the possibilities obtained directly following insertion of the mini-Tn*5 TF* into the genome. The ends of the mini-transposons are simply described as I for Iend and O for Oend. The two crosses (in green) show that the case does not lead to any potential deletion but to inversion of the genomic fragment. For a better understanding of the table the mini-Tn*5 TF* is shown as a blue box (represented as a *FRT* site without the rest of the mini-transposon), mini-Tn*5 KpF* as a green box and the new DNA junction as a red box, in case 5 to 28. These cases represent the 24 different combinations after the second insertion of mini-Tn*5 KpF*.

#### 4.2.2.1.2 Optimizing the genome reduction: blind random deletion

In order to improve the technique for random deletion, a step once considered to be important until recently was removed from the procedure. Generation of the single and double mutant libraries during the establishment of the first *P. putida* $\Delta_1$ mutant was here suppressed from the experiments. All the steps, from mini-Tn*5* derivatives insertion in the genome to pBBFLP insertion in the exconjugants, were carried out without any intermediate (e.g. streaking and screening), thus avoiding the loss of potential candidates for deletion.

*P. putida* $\Delta_1$-91 and *P. putida* $\Delta_1$-407 mutants were separately used as new wild-type strains for the insertion of both mini-Tn*5* derivatives. In both cases mini-Tn*5 KpF* was first inserted by triparental mating and the resulting *P. putida* $\Delta_1$ SMT mutants were processed as described in section 3.5.

Out of the three *FRT* sites present in the genome after all conjugation steps, the position of only a single one was known. All the different possible combinations are summarized in Table 4-1. Twenty-four cases are summarized all with the same possibility of occurrence due to the random aspect of the transposition.

For all the 24 possible combinations of mini-Tn*5* derivatives and new DNA junction, the corresponding phenotypes mentioned hereafter are always considered after a possible genomic rearrangement, meaning that no conditionally essential genes are situated in the target genomic fragment. An overview of the possible phenotypes for the double-deleted mutants in each of the recombination case is given in Table 4-2.

**Table 4-2: Resulting phenotype for each possible *P. putida* $\Delta_2$ mutants.**

| Phenotype / Case | $Km^S$ and $Tel^S$ | $Km^S$ and $Tel^R$ | $Km^R$ and $Tel^S$ | $Km^R$ and $Tel^R$ |
|---|---|---|---|---|
| 5, 7 | | | | |
| 6, 15, 16, 19 | | | | |
| 8, 10, 13, 17, 20, 26, 28 | | | | |
| 9, 22, 25 | | | | |
| 11, 12, 14, 18, 23, 24, 27 | | | | |
| 21 | | | | |

As mentioned in paragraph 4.2.2.1.1.2, out of the 24 combinations only seven lead to genomic deletion with the loss of both resistance cassettes (second column of Table 4-2). In cases 5 and 7 (second row), three different options lead to sensitivity either to Km, to Tel or to both in parallel (green cells). In case 21 the options are reduced to $Tel^S$ or $Km^S/Tel^S$ in parallel (green cells of row 7). For the rest of Table 4-1 four cases (6, 15, 16 and 19) can lead to Km sensitivity (green cell in the 3rd row) and three cases (9, 22 and 25) to the loss of Tel resistance (green cell in the 5th row). Finally seven combinations do not offer any option to lose the resistances (green cell in 6th row).

Using this procedure, four independent double mutants, two *P. putida* $\Delta_2$-91 and two *P. putida* $\Delta_2$-407 were found. The procedure was faster than the one previously described and more efficient to provide mutants. The double-deleted mutants, named *P. putida* $\Delta_2$ followed by the name of the first deleted mutant, will be described in the next two paragraphs. The determination of the position of the new mini-Tn*5* derivatives will be also explained as no mapping was carried out after each insertion.

### 4.2.2.2 *P. putida* $\Delta_2$-91 knockout mutants

#### 4.2.2.2.1 Phenotypic confirmation

In total 310 colonies were tested for sensitivity to Km, Tel and Pip and for resistance to FOA after action of the Flp recombinase. Four colonies did not grow on medium containing uracil and citrate as carbon source; therefore were not considered for further analyses. Incubation of the piperacillin containing plate showed that 87 colonies were sensitive to the compound meaning that in 28.4 % of the cases a true transposition event occurred. Among these 87 colonies eight were resistant to Km but all grew on Tel medium. Even though both resistances were not lost, the sensitivity to kanamycin indicated a genomic rearrangement. Five putative mutants with the $Km^S/Tel^R$ phenotype were selected out of the eight for further experiments.

The apparent phenotype decreased the amount of different possible combinations for the positions of the mini-Tn*5* derivatives before deletion. From 24 the number was reduced to five possible cases (5, 6, 15, 16 and 19) after phenotypic analyses, see the 3rd column from Table 4-2. In order to verify these combinations and to reduce the options, PCR amplifications were carried out.

#### 4.2.2.2.2 Confirmation by PCR amplifications and sequencing

The involvement of the *FRT* site, from the new DNA junction of *P. putida* $\Delta_1$-91 mutant, with the *FRT* of one of the new mini-Tn*5* derivatives inserted or the loss of this *FRT* site was verified by PCR. The primers PP_3490 F and PP_3529 R were used to amplify the new DNA junction as it was previously done for the verification of the genomic deletion in *P. putida* $\Delta_1$-91 strain. As positive control the gDNA of the single-deleted mutant was taken, as well as the gDNA of a colony which after the second deletion harbored a $Km^R/Tel^R$ phenotype. The gDNA of *P. putida* TEC1 and of *P. putida* T91 (before the first deletion) was used as negative control.

The results of the amplification are presented in Figure 4-24.

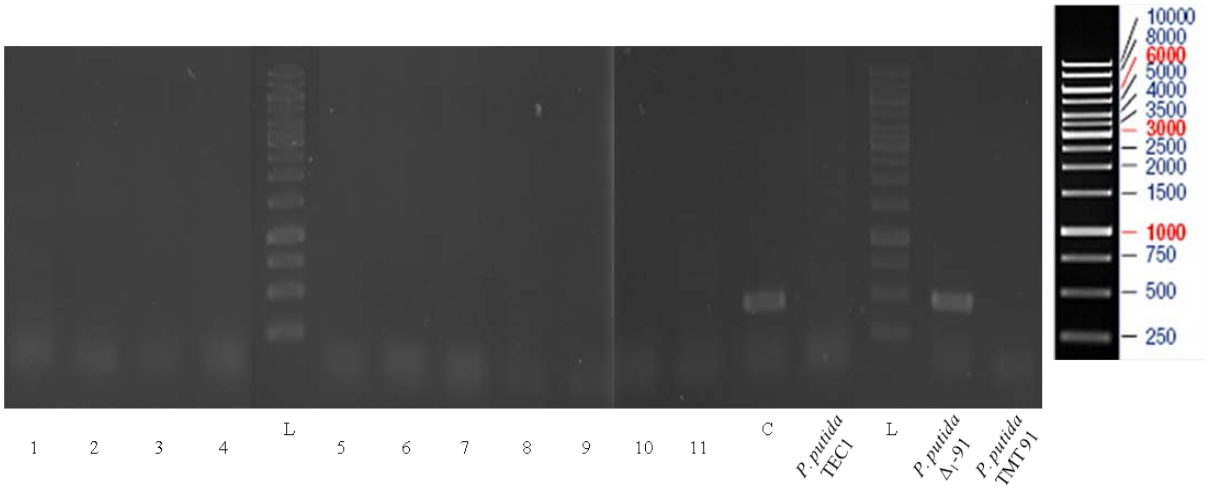



**Figure 4-24: PCR verification of the participation of the *FRT* site from the new DNA junction in the second deletion in putative *P. putida* $\Delta_2$-91 mutants.**
Gel electrophoresis with the putative *P. putida* $\Delta_2$-91 mutants (1 to 11) used for the PCR with primers PP_3490 F and PP_3529 R. Four control strains are loaded: *P. putida* $\Delta_1$-91 TMT (after deletion step, $Km^R/Tel^R$ phenotype (C) and *P. putida* $\Delta_1$-91 as positive controls, *P. putida* TEC1 and *P. putida* TMT 91, as negative controls. The ladder (L) is loaded twice and the size of the bands are shown on the right side. Two signals appear over the gel corresponding to the new DNA junction size in the single-deleted mutant and the failed second deleted mutant.

As can be seen in Figure 4-24, a signal with the expected size of ~400 bp (section 4.2.1.3.1) only appeared for both positive controls. No signal was detected for *P. putida* TEC1 and *P. putida* TMT 91 which confirmed them as negative controls. For the putative double-deleted mutants (1 to 11) no signal was found on the gel meaning either that the new DNA junction was framed by both mini-Tn*5*, in *P. putida* $\Delta_1$-91 TMT mutants, and the recombination occurred between them or that one *FRT* site from a mini-Tn*5* derivative recombined with the *FRT* site from the new DNA junction. In this last case, the apparent phenotype of the putative mutants involves the recombination between the mini-Tn*5 KpF* and the *FRT* site from the new DNA junction due to the loss of kanamycin but not of tellurite resistance. These results confirmed the hypothesis of the five possible configurations of *FRT* sites before the second deletion step, cases 5, 6, 15 and 16.

In order to characterize precisely the second deletion and to understand better how the deletion occurred, mutants were submitted to AP-PCR and sequencing. In total four colonies were chosen for the detection of the second excision, but only the position of the mini-Tn*5 TF*, which did not participate to the deletion, was determined. In three cases gene PP_1626 (*mutS*) was disrupted. Thus the colonies seemed to originate from the same *P. putida* $\Delta_1$-91 TMT mutant. The negative orientation and the position in the chromosome of the mini-Tn*5 TF* only matches with case 16. In the last case gene PP_0659 (*metB*) was disrupted. The position and orientation of the mini-Tn*5 TF* inserted in this case indicated that case 15 was the only possible combination. Several AP-PCR involving different

sets of primers were tested for discovering the deletion between mini-Tn*5 KpF* and the *FRT* site from the new DNA junction, but none of them were successful. The scheme of the deletions is shown in Figure 4-25.

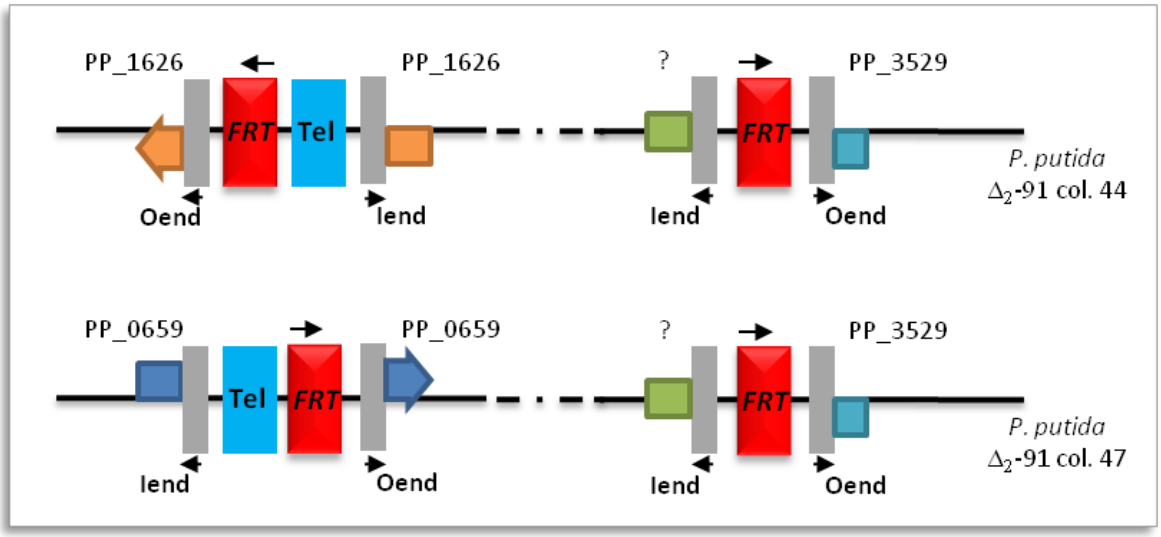



**Figure 4-25: *FRT* combinations in *P. putida* $\Delta_2$-91 mutants.**
Both double knockout mutants (col. 44 and 47) carry two *FRT* sites (red box) in the genome after deletion. Col. 44, $Km^S/Tel^R$, corresponds to the deletion occurred in a *P. putida* $\Delta_1$-91 TMT strain in the same configuration as case 16, Table 4-1. The position of the mini-Tn*5 TF* was detected (PP_1626). Col. 47, $Km^S/Tel^R$, corresponds to the deletion occurred in a *P. putida* $\Delta_1$-91 TMT strain in the same configuration as case 15, Table 4-1. The position of the mini-Tn*5 TF* was detected (PP_0659). The flanking disrupted region is represented by green rectangles when the position was not mapped by AP-PCR and sequencing.

The action of the flippase on *P. putida* $\Delta_1$-91 TMT generated two $\Delta_2$-91 mutants. Both of them carry the tellurite resistance marker but have lost the kanamycin cassette as well as the wild-tape *pyrF* operon. Mini-Tn*5 TF* insertion was mapped in both cases but the exact determination of the deleted genes was not yet successful.

Originating from *P. putida* $\Delta_1$-91, the double-deleted mutants lack the *ilvE* gene. However, it was interesting to assess the influence of the second deletion on the growth of the different mutants in M9 medium. Growth of the mutants was measured using the Bioscreen C MBR (Oy Growth Curves Ab Ltd, Helsinki, Finland; section 3.6.2). Preliminary results revealed that one of the double mutants (col. 44) was able to grow without valine and isoleucine when citrate or succinate was supplemented as a carbon source. In the presence of the amino acids the three mutants tend to show the same behavior. However, the growth experiments were not sufficient to obtain a clear conclusion regarding the growth of col. 44 in both media without the supplementation of one of the branched chain amino acids.

#### 4.2.2.3 *P. putida* $\Delta_2$-407 knockout mutants

The successive insertions of both mini-Tn*5* in the genome of *P. putida* $\Delta_1$-407 led to the generation of *P. putida* $\Delta_1$-407 TMT mutants.

##### 4.2.2.3.1 Phenotypic confirmation

In total 200 colonies were tested for sensitivity to Km, Tel and Pip and for resistance to FOA after action of the Flp recombinase. All the streaked colonies grew on the basic medium supplemented with uracil and citrate as carbon source. Incubation of the piperacillin containing plate showed that 125 colonies were sensitive to the compound meaning that in 62.5 % of the cases a true transposition event occurred. Among these 125 colonies 56 were found to be sensitive to Km and Tel simultaneously and 54 to have lost only the resistance to kanamycin ($Km^S/Tel^R/Pip^S$). Five putative mutants with the $Km^S/Tel^R$ phenotype and five with $Km^S/Tel^S$ phenotype were selected for further experiments.

For the putative $\Delta_2$-407 mutants the situation is different as previously mentioned due to the presence of two different phenotypes. In the case of sensitivity to kanamycin only, five possible combinations for the *FRT* sites were highlighted (cases 5, 6, 15, 16 and 19) in paragraph 4.2.2.2.1. However, the sensitivity to both Km and Tel indicate different possible combinations for the *FRT* sites before deletion. In total eight cases are concerned: 5, 7, 8, 13, 17, 20, 26 and 28 from Table 4-1. These eight cases can be categorized in two groups depending on the amount of *FRT* sites left in the genome after deletion. In cases 5, 8, 13, 17 and 28 the *FRT* sites interacting involve the complete loss of one of them and the recombination of the two other, leading to the presence of a single recognition site in the double knockout mutant. In cases 20 and 26 only the *FRT* sites from the newly inserted mini-transposons derivatives were involved in the action of the flippase, leading to the presence of two recognition targets in the double knockout mutant. In the case 7 both options are possible. In order to verify these combinations and to reduce the options, PCR amplifications were carried out.

##### 4.2.2.3.2 Confirmation by PCR amplifications and sequencing

In order to verify the phenotype of all the putative mutants, two different PCR amplifications were carried out. Two different targets were searched for: <Km-*pyrF*> cassette (amplification with primers Km-pBAM F/Iend Rev2, see Tables 8-2 and 8-3) and tellurite cassette (primers Tel fw/Tel rev, see Table 8-3). No signal was detected on the gel for the <Km-*pyrF*> cassette, which confirmed the loss of the resistance and signal for the tellurite cassette was exclusively found for the colonies resistant to tellurite on the plates.

Two colonies, *P. putida* $\Delta_2$-407 col. 1 and col. 3, one from each phenotype, were randomly chosen among the ten tested by PCR and AP-PCR was carried out. Different settings were applied in order to adapt to the possible *FRT* combinations and resulting deletion in the two double knockout mutants.

With the different sets of primers used, it was possible to determine which recombination case precisely happened for each of the mutants.

- *P. putida* $\Delta_2$-407 col. 1 originated from a *P. putida* $\Delta_1$-407 TMT with the same configuration as represented in case 7, Table 4-1. Recombination occurred between the mini-Tn*5 KpF* and mini-Tn*5 TF* newly inserted in the genome of *P. putida* $\Delta_1$-407, meaning that after deletion two *FRT* sites were present in the genome of *P. putida* $\Delta_2$-407 col. 1. The result of the deletion is shown in Figure 4-26.
- *P. putida* $\Delta_2$-407 col. 3 originated from a *P. putida* $\Delta_1$-407 TMT with the same configuration as represented in case 16, Table 4-1. Recombination occurred between the *FRT* site of newly inserted mini-T*n5 KpF* and the *FRT* site from the new DNA junction of *P. putida* $\Delta_1$-407, meaning that after deletion two *FRT* sites were present in the genome of *P. putida* $\Delta_2$-407 col. 3. The result of the deletion in shown in Figure 4-26.

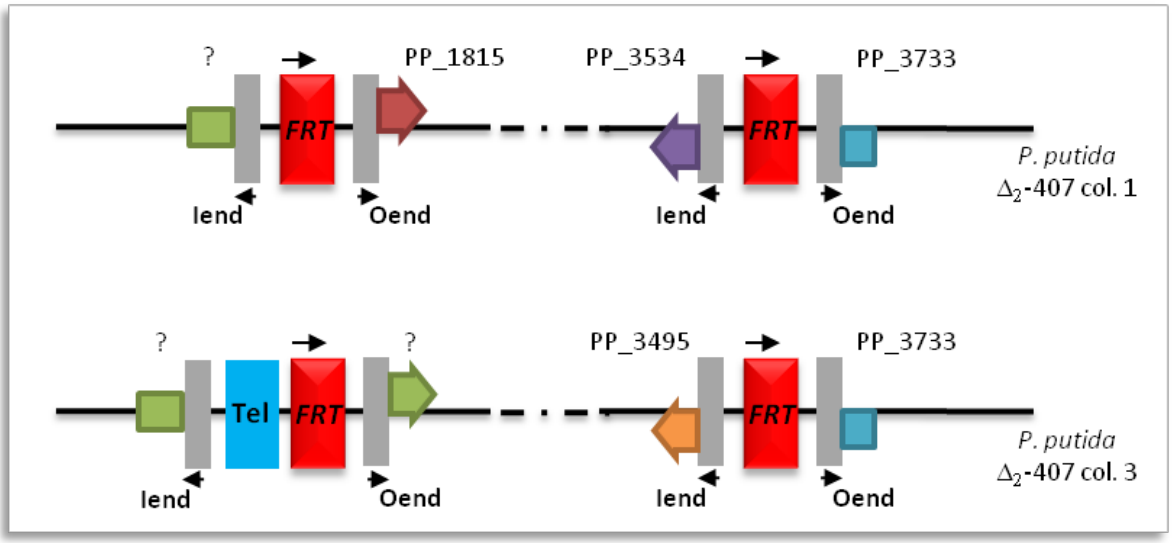



**Figure 4-26: *FRT* combinations in *P. putida* $\Delta_2$-407 mutants.**
Both double knockout mutants (col. 1 and 3) carry two *FRT* sites (red box) in the genome after deletion. Col. 1, $Km^S/Tel^S$, carries the newly formed DNA junction appeared in *P. putida* $\Delta_1$-407, and a new DNA junction. One flanking gene (PP_1815) was detected. Col. 3, $Km^S/Tel^R$, carries a newly formed DNA junction, FRT site of mini-Tn*5 KpF* (inserted in PP_3495) recombined with the *FRT* site left in the genome of *P. putida* $\Delta_1$-407, and the mini-Tn*5 TF* which did not interact after its insertion in the genome. The flanking disrupted regions are represented by green rectangles when the position was not mapped by AP-PCR and sequencing.

Exact position of the deleted genomic fragment was determined only for *P. putida* $\Delta_2$-407 col. 3. Interestingly, the second deletion occurred upstream the first excision; therefore comprises part of the deletion from *P. putida* $\Delta_1$-91 (deleted between PP_3490 and PP_3529). The phenotype of the other recombinant mutants indicated the presence of genomic rearrangement corresponding theoretically to the second deletion of a genomic fragment, but it could not be mapped until now. The next step consisted on the assessment of the behavior and new characteristics of the double-deleted mutants.

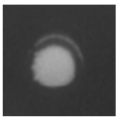
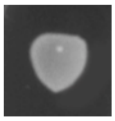
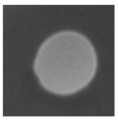
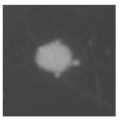
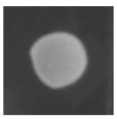
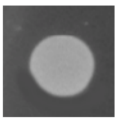
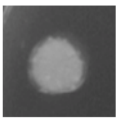
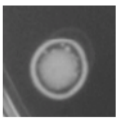
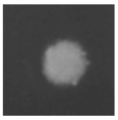
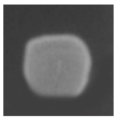
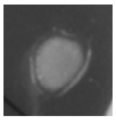
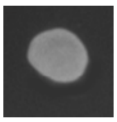
## 4.3 Assessment of the characteristics of all deleted mutants and comparison with the *Pseudomonas putida* TEC1 wild-type strain

### 4.3.1 Screening of the mutants

The loss of gene PP_3587 (*tpx*), coding for a redoxin domain protein (Tpx, Thiol peroxidase) and gene PP_3668, coding for a catalase/peroxidase HPI, in the *P. putida* $\Delta_1$-407 mutant and thus in the two *P. putida* $\Delta_2$-407, indicated independently a possible attenuation in the response to oxidative stress. Oxidative stress can be generated by the presence of oxygen derivatives such as superoxide and hydrogen peroxide. To assess the consequences of this loss, the mutants were grown with different concentrations of hydrogen peroxide ($H_2O_2$) as described in paragraph 3.6.1. All the selected mutants participated to the test including also the *P. putida* $\Delta_x$-91 series. As control *P. putida* TEC1 wild-type strain was utilized.

One LB plate, corresponding to one $H_2O_2$ concentration, was used to test all the different strains. The results of the oxidative stress are shown in Table 4-3.

**Table 4-3: Oxidative stress experiments with the different deleted mutants in comparison with the *P. putida* TEC1 wild-type strain.**
Out of the nine different concentrations used for hydrogen peroxide in the medium only two allowed some of the strains to grow (2 and 4 mM). The first column indicates the plates which were supplemented with uracil and citrate but without hydrogen peroxide (control). One drop of each bacterial culture was used on each plate. The corresponding growth is reported in the Table.

| $H_2O_2$ concentration / Strains | Ø | 2 mM | 4 mM |
|---|---|---|---|
| *P. putida* TEC1 |  |  |  |
| *P. putida* $\Delta_1$-91 |  |  |  |
| *P. putida* $\Delta_2$-91 col. 44 |  |  |  |
| *P. putida* $\Delta_2$-91 col. 47 |  |  |  |
| *P. putida* $\Delta_1$-407 |  |  |  |
| *P. putida* $\Delta_2$-407 col. 1 |  |  |  |
| *P. putida* $\Delta_2$-407 col. 3 |  |  |  |

As can be seen in Table 4-3, all the deleted *P. putida* $\Delta_x$ mutants grew on LB agar supplemented with citrate and uracil when no oxidative stress was applied (first column). On medium supplemented with hydrogen peroxide in a concentration comprised between 2 and 4 mM the control strain grew as well as the deleted mutants from the *P. putida* $\Delta_x$-91 series. Morphology of the drop is considered as being the same when 2 mM are added and slightly modified when the concentration was increased to 4 mM representing a decrease in the growth. On the contrary already with 2 mM *P. putida* $\Delta_1$-407 and *P. putida* $\Delta_2$-407 col. 1 had a reduced growth and nothing was detected for *P. putida* $\Delta_2$-407 col. 3. Furthermore when $H_2O_2$ was added to a final concentration of 4 mM, none of the *P. putida* Δx-407 mutant series survived. From 8 to 250 mM growth was detected neither for the wild-type nor for the deleted mutant strains. This modification in the ability to grow matches with the loss, only in the 407 mutant series, of the PP_3587 and/or PP_3668 gene.

### 4.3.2 Growth comparison of the mutants and wild-type strain

Genetic, phenotypic and stress response analyses indicated obvious differences between all the mutants generated during the PhD thesis. In order to proceed with the last comparison of the single and double-deleted mutants with the wild-type *P. putida* TEC1 and among each other, growth tests were carried out using the Bioscreen C MBR. M9 minimal medium was used as the common basic medium and uracil was supplied independently of the strain inoculated and the final composition of the medium used. Citrate, glucose, succinate, glycerol and benzoate were the different compounds used separately as carbon source in the medium. Besides these five sources a last test was carried out without any carbon delivery as explained in paragraph 3.6.2. Due to the loss of *ilvE* in *P. putida* $\Delta_x$-91 mutant series and their consequent auxotrophy for branched-chain amino acids, valine and traces of isoleucine were supplied in the medium. Due to the loss of *argC* in *P. putida* $\Delta_x$-407 mutant series and a possible participation in the affection of the growth of the strains, arginine was supplied as well but separately. In total different graphics were established summarizing the growth of each mutant in a given medium and comparing the growth of a single strain in the different media (data not shown).

#### 4.3.2.1 General overview of the growth curves

Growth experiments were carried out by duplicates. After building the curves, the duplicates revealed high similarity for each mutant in each medium. Due to the use of *P. putida* TEC1 as wild-type for the generation of the deletion, each medium was supplemented with uracil in order to satisfy the auxotrophy of the strains. Test of the bacterial growth in M9 minimal medium supplemented only with uracil did not allow any strains to grow (OD values constant around the initial OD of inoculation), showing the absence of any carbon or energy sources.

The first highlight of the different sets of curves concerns the dominance of the *P. putida* TEC1 growth. In 15 of the 17 media (M9+ura is not considered here) used for the growth experiments, the

wild-type strain reached the highest optical density. This confirms the general idea that when a strain loses parts of its genome it becomes gradually unstable or the fitness is decreasing with the importance of the undergone deletion. The term “importance” involves here the size of the excision and the number of non essential genes involved in metabolic pathways or maintain of the cell, etc…. However, in M9 medium supplemented with uracil and citrate or succinate, *P. putida* TEC1 grew until a lower OD than some mutant strains grew, and in the case of succinate with a longer lag phase (12 h before the start of the exponential phase instead of 4 to 5 h for the faster and better growing strains). In more detail, supplementation with citrate allows *P. putida* $\Delta_2$-407 col. 1 to grow until $OD_{max}$ = 0.427 directly superior to the wild-type with an $OD_{max}$ = 0.355. In the case of succinate three strains, *P. putida* $\Delta_2$-91 col. 44, *P. putida* $\Delta_1$-407 and *P. putida* $\Delta_2$-407 col. 1, exhibited higher $OD_{max}$ than the wild-type ($OD_{max}$ = 0.261). Interestingly, in these two specific substrates the addition of any amino acid inverts the tendency and allows the wild-type strain to improve its grows by 1.7 to 2,2-fold in the case of supplementation with arginine and by 3.1 to 5-fold with valine and traces of isoleucine. On the contrary the mutants with better growth did not seem to be affected by the addition of the amino acids in the medium as the general tendency of the curves stay the same.

In the case of glucose as substrate, the supplementation of amino acid seems to affect also the growth of *P. putida* TEC1 (increase by 1.3-fold with arginine and 2.1-fold with valine and isoleucine); however, this improvement in the growth does not compete with any of the mutants as none of them ever overtook the growth of the wild-type.

Addition of valine and isoleucine affects the growth of *P. putida* $\Delta_x$-91 mutant series in the case of glycerol-based medium. In this case only a clear increase of the growth appears on the graphs, almost 7-fold more than in glycerol without amino acid and 3.6-fold more than in the medium supplemented with arginine.

# 5. Discussion

*Pseudomonas putida* strain KT2440 is a ubiquitous organism with a considerable metabolic diversity and efficient stress response. Different mutants have been generated for both strains. Among them, *P. putida* TEC1, carries a partial deletion of the *pyrF* gene which causes an auxotrophy for uracil, it is considered to endow the same metabolic properties as the original strains (Galvao and de Lorenzo, 2005). This feature allowed efficient genetic manipulation of the organism for genomic deletion using the *pyrF* gene as either a selectable or counterselectable marker.

A major challenge in Synthetic Biology is the generation of a stable, streamlined and yet robust microbial chassis in which the metabolic and regulatory wiring is well characterized and that can be subsequently used for rational cellular re-programming. Two different approaches have been followed to tackle this challenge. The "bottom-up" approach aims to build simple microbes either from their basic components (e.g. nucleic acids, proteins (Chiarabelli *et al.*, 2009)) or by first synthesizing a genome from scratch and then booting it up into a suitable chassis (Gibson *et al.*). The "top-down" approach in contrast aims to streamline, modify and control the different units and modules of an existing organism. By removing non-essential and unnecessary genes for sustaining a living cell, this approach has also contributed to the study of the concept of a minimal cell and of the principles underlying the origin of life. Various organisms have been used as a template for gene removal and assessment of gene essentiality, e.g., *M. genitalium*, *H. influenza* and *E. coli* (Hutchison III *et al.*, 1999; Hashimoto *et al.*, 2005; Pósfai *et al.*, 2006).

In dissertation work, the genome of *P. putida* TEC1 was streamlined through gene deletions by using an approach based on the semi-random excision of groups of genes via the combination of different mini-transposons and the site-specific recombination system. The main advantage of this method is that it can be used without prior knowledge of which genes are dispensable. The global goal of this study is to "simplify", better understand and, ultimately, control the wild-type strain. "Deconstruction" of bacteria by reducing their genomes will allow their manipulation by re-organizing the strains with new specific genetic circuits. This method will be, in the short term, directly adapted to *P. putida* strain KT2440.

## 5.1 A combinatorial deletion method adapted to *Pseudomonas putida*

### 5.1.1 Mini-transposon mutagenesis and site-specific recombination system

The development of a series of tools and application to *P. putida* TEC1 strain allowed the generation of two consecutive reductions of the genome. Random mutagenesis associated with site-specific recombination created two independent single-deleted mutants (*P. putida* $\Delta_1$-91 and $\Delta_1$-407) and four

corresponding double-deleted strains (*P. putida* $\Delta_2$-91 col. 44 and 47 and $\Delta_2$-407 col. 1 and 3). The deletion of parts of the genome was characterized genetically during the work of this dissertation work, see sections 4.2.1.3 and 4.2.1.4.

For this study, two mini-transposons derivatives, mini-T*5 KpF* and mini-Tn*5 TF*, were used for the genomic deletion procedure. The singularity of these transposable elements is based on their divergent selectable and counterselectable markers. The first mini-transposon derivative comprised two different traits: (i) the *pyrF* operon, which confers the ability to synthesize uracil and yields a FOA-sensitive phenotype and (ii) the resistance to kanamycin which, associated with the presence of the *pyrF* operon, allows efficient screening of the *P. putida* SMT and *P. putida* $\Delta_x$ SMT mutants. In the past, the particular property of the *pyrF* gene, being used as a selectable (uracil prototrophy) and a counterselectable (FOA sensitivity and toxicity for the cell) marker, has been efficiently exploited in different bacteria such as *E. coli* and *Mycobacterium smegmati*s (Husson *et al.*, 1990; Knipfer *et al.*, 1997). The second mini-transposon used in this work provided the resistance to potassium tellurite which is usually toxic for the wild-type strain. This second marker simplifies the recognition of the *P. putida* TMT, *P. putida* $\Delta_x$ and $\Delta_x$ TMT mutant strains. The transposition event is directly visualized on the plates by the production of a characteristic black color due to the transformation of potassium tellurite into tellurium (Sanchez-Romero *et al.*, 1998).

Theoretically, the loss of the genomic fragments involved that the mutant loses its resistance markers and thus, do not play any further role in the subsequent experiments. This is an advantage of the method as compared to other techniques in which knockout mutant strains are generated carrying antibiotic resistance markers and are further released in the environment or used for live vaccines (Nakayama *et al.*, 1988; Herrero *et al.*, 1990). The use of a non-antibiotic resistance marker in this study was an alternative to the kanamycin resistance in order to efficiently select the *P. putida* TMT and further $\Delta_x$ TMT mutants.

#### 5.1.1.1 An alternative use for mini-transposon insertion

With the development of the sequencing techniques over 1,300 fully sequenced genomes are currently available (http://www.ncbi.nlm.nih.gov/sites/genome, http://www.genomesonline.org/cgi-bin/GOLD/bin/gold.cgi?page_requested=Complete+Published). To increase our understanding of the actual gene functions, regulatory networks and general behavior of the organisms, mutagenesis experiments are routinely carried out under both laboratory and environmental conditions. One common strategy is to use mini-transposons, which are randomly inserted into chromosomes or plasmids. Several mutant libraries were constructed in the past with Eukaryotes (Ross-Macdonald *et al.*, 1999) and Prokaryotes (Gerdes *et al.*, 2002; Lehoux *et al.*, 2002; Jacobs *et al.*, 2003), as mentioned in section 1.3.3. Each of them has generated several thousands of mutants.

***Selection and establishment of mini-Tn5 mutant libraries***

Here, a library of 613 *P. putida* TMT mutants carrying two mini-transposons was established in M9 minimal medium in order to subsequently generate deleted mutants in which unstable or unknown and non-essential elements are excised under these conditions. The first 61 *P. putida* SMT mutants were successfully mapped (Figure 4-5) and ten candidates were chosen for the following mini-Tn*5 TF* insertion and generation of the 613 *P. putida* TMT mutants. The genomic fragment to potentially delete was situated between both mini-transposons. Molina-Henares and colleagues isolated in total 7,760 independent mini-Tn*5* mutants representing a genome-wide library; however, they were generated on LB medium and only further tested on M9 medium (Molina-Henares *et al.*, 2010). Most of the disrupted genes obtained here were already studied in LB and M9 media. The potential candidates for genomic fragment deletion were pre-selected in function of the combination of both mini-Tn*5* derivatives in their genome, involving the orientation and the "distance" between both of them. The analysis of the predicted knockout genes and gene functions, located between both mini-Tn*5* derivatives, did not play a role in the pre-selection of the candidates as their simultaneous deletion might render some of them essential under given conditions. Increasing the number of gathered colonies *per se* is not crucial to improve the method, but increasing the number of mapped mini-Tn*5* within the library would certainly provide more potential candidates for genomic deletion.

For the generation of deleted mutants used for diverse biocatalytic processes some of the removed genes might become essential in M9 supplemented with other carbon source. In the context of this study, growth of the putative TMT mutants on M9 supplemented with citrate, directly after transposition, already implies the reduction of the number of potential mutants due to the random knocking out of some essential genes under these conditions. Even though *P. putida* $\Delta_1$-407 could be obtained directly in M9, *P. putida* $\Delta_1$-91 was first selected in LB, due to the presence of the later revealed *ilvE* essential gene, and subsequently tested in M9. The libraries of mini-transposon insertion mutants that created several thousands of mutants for *P. aeruginosa* and *P. putida* were established in LB medium and the detailed phenotypic analyses of the mutants were carried out afterwards (Jacobs *et al.*, 2003; Duque *et al.*, 2007). It is then understandable that with the stringent growth conditions used here a lower amount of mini-Tn*5* inserted mutants were directly collected after transposition. Finally, M9 minimal medium has the added advantage of counterselecting *E. coli* strains. When mutants had to be grown in LB, the counterselection was then possible by the supplementation with Nalidixic acid (Goss *et al.*, 1965).

***Randomness***

The randomness of the transposition event was established in several studies, considered from the point of view of the integration sites (Stachel *et al.*, 1985; Berg, 1989; De Lorenzo *et al.*, 1990; Duque *et al.*, 2007; Puttamreddy *et al.*, 2010). Nevertheless, some publications evoked the existence of hot

spots in which a transposable element may be either inserted frequently (warm spots) or infrequently (cold spots) (Davies and Hutchison, 1995; Gerdes *et al.*, 2002). Here, different aspects were verified for the mutant libraries generated. First, 88.5 % and 90.4 % of the SMT and TMT (mini-Tn*5 TF* taken into account) mutants, respectively, carry a mini-Tn*5* derivative in a different position of the genome. These values agree with the results found for the genome-wide mutant library of *P. putida* KT2440 (Duque *et al.*, 2007), in which the random transposition corresponded to 95 %. Secondly, considering both SMT and TMT mutant libraries and based on Figures 4-5 and 4-7, it was shown that the mini-Tn*5* derivative targets genes independently from the template strand for transcription. The transcription of 23 (45 %) and 127 (55 %) hit genes uses the lagging strand as template in the SMT and TMT mutants, respectively. Thus, the orientation of the ORFs does not seem to influence the insertion event of the mini-Tn*5* derivative. It was shown that putative promoter sequences can be found in the mini-transposons with a low but sufficient level of transcription that might be used by the cells. These sequences were not intentionally inserted, but arose possibly by the combination of the different fragments during the construction of the mini-Tn*5* derivatives (Gerdes *et al.*, 2002). Thirdly, there is no evidence for the automatic insertion of the mini-Tn*5* derivatives into a region of the genome with a specific value of the GC content (Figures 4-5 and 4-7). The last aspect of randomness that must be considered is the influence of the position and direction of the first inserted mini-transposon on the second transposition. Considering the 255 mapped *P. putida* TMT, 116 mutants (45.4 %) exhibited the mini-Tn*5 KpF* and mini-Tn*5 TF* in the same direction and 63 carried the mini-transposons in the correct order (Figure 4-9, case C) entailing the loss of the resistance cassettes after the deletion step. All these results confirm the expected proportions for each case of recombination. The direction of the mini-Tn*5 KpF* already present in the genome did not seem to influence the way of insertion of the mini-Tn*5 TF*, as previously expected.

In summary, the transposition event with respect to the place of insertion and direction of insertion was random, confirming that previously described by Shevchenko and colleagues (Shevchenko *et al.*, 2002). The presence of specific hot spots, however, was more difficult to prove due to the low number of selected mutants. Nevertheless, by increasing the library of SMT and TMT mutants, it would be possible to make such an assessment in the future.

#### 5.1.1.2 Adaptation of the Flp-*FRT* site-specific recombination system

***Applications of the Flp-FRT system***

In the current work, the Flp-*FRT* recombination system played a key role in the deletion of the genomic fragments in *P. putida*. This method is the most adaptable to the *in vivo* excision system and is used in a broad host-range of bacteria, e.g. *E. coli* and *P. aeruginosa* (Huang *et al.*, 1997; Hoang *et al.*, 1998), yeasts, e.g. *S. cerevisiae* (Storici *et al.*, 1999) and plants, e.g. tobacco (Gidoni *et al.*, 2001). Moreover, the action of the recombinase was shown to be extremely efficient for removing small and

larger fragments (Cherepanov and Wackernagel, 1995; Datsenko and Wanner, 2000). In the latter case, a 5,178-nt segment was successfully excised from the genome of an *E. coli* strain (corresponding to the *lacZYA* operon) without affecting cell survival. For this purpose, the fragment to be removed was replaced by a resistance cassette flanked by two directly repeated *FRT* sites. The resistance was excised from the genome by action of the flippase on the two recognition targets. In the present study the principle of removing the flanked fragment remained the same; however, the two *FRT* sites were inserted randomly and independently from each other.

In *E. coli* and *Pseudomonas* species, the method has been differently applied for the rescue of pathogenicity islands (Posfai *et al.*, 1997), integration and expression of elements into specific chromosomal locations (Hoang *et al.*, 2000) and site-specific insertion of heterologous DNA into the chromosome (Huang *et al.*, 1991, 1997). Pósfai and colleagues adapted the recombination system for the deletion of the pathogenicity island LEE of an *E. coli* strain. Two sets of small suicide plasmids were constructed, each of them were carrying a single *FRT* site and a homologous sequence to the target sites of the genes to be removed. After homologous recombination of each plasmid in the genome, the flippase was expressed leading to the deletion of the island located between both *FRT* sites. The fragment to be removed was targeted prior to the experiments and specific plasmids were designed in order to integrate the *FRT* sites into the chromosome. This method is similar to that herein developed but it diverges in the pre-selection of the genes to be deleted (Posfai *et al.*, 1997). Huang and colleagues established a new method called FLIRT (Flp-mediated DNA integration and rearrangement at prearranged genomic targets). As a first step, the *FRT* sites in association with transposable elements were randomly inserted into the genome of *E. coli* for the successive targeting of exogenous DNA in those sites. The association of *FRT* sites with transposable elements corresponds to the idea developed here, where the insertion of the recognition sites into a chromosome is performed randomly. In 2005, Kato & Hashimoto generated new deletions in the genome of *E. coli* K12-MG1655 using the Flp-*FRT* recombination system (FLP-FRT2) in which two suicide plasmids were engineered with a single *FRT* site in each to further transfer the excised fragment to the mini-F plasmid (Timmis *et al.*, 1975; Kato and Hashimoto, 2007). This rescued fragment was also employed to detect the presence of *cis*-acting chromosome regions in those regions that were not deleted in a previous study (Hashimoto *et al.*, 2005).

***Different site-specific recombination systems***

The combination of mini-transposons with a site-specific recombination system allows random genomic deletions influenced by the cell itself. The pool of different *P. putida* TMT strains carrying two mini-Tn*5* derivatives and thus, two *FRT* sites, provides many possibilities for the cells to excise unnecessary genomic fragments by the action of the flippase under specific conditions. Yu and colleagues designed a similar experiment creating a method for genomic deletions in *E. coli* strain K12-MG1655 (Blattner *et al.*, 1997; Yu *et al.*, 2002). The system was based on two Tn*5* transposons

combined each with a single *loxP* recombination site, generating the two derivatives TnKloxP and TnCloxP. The Cre/*loxP* recombination system of bacteriophage P1, developed by B. Sauer, interacts in a similar way as the Flp-*FRT* system (Sauer, 1987; Jayaraman *et al.*, 2002). Each transposon derivative was inserted separately in different cells of *E. coli* and the mapping was performed by sequencing. In this case two parallel mutant libraries were obtained, regrouping 800 strains. Phage P1 transduction was used to combine two mutations into the same strain (one from TnKloxP and one from TnCloxP) depending on the genes to be deleted. The pre-selection was based on the position and direction of the *loxP* sites and the presence of potentially essential genes. The same conditions were required for the pre-selection of *P. putida* TMT strains for deletion in this study (Figure 4-9). The possible presence of essential genes was assessed with the assistance of the two genome-scale metabolic reconstructions established for *P. putida* KT2440 and the genome-wide mini-Tn*5* mutant library (Nogales *et al.*, 2008; Puchalka *et al.*, 2008; Molina-Henares *et al.*, 2010). However, as previously mentioned the effect of two simultaneous combined genes cannot be predicted with accuracy. Thus, the pre-selection was here based on the position of the *FRT* sites.

***Combination of transposon mutagenesis and site-specific recombination system***

Using the combined methods of transposon insertion and the Cre/*loxP* system, Yu and colleagues were able to generate six mutants of *E. coli* strain K12-MG1655, deleting in total, 472 genes (10.9 % of the whole genome). The single-deletions ranged between 1.27 and 2.5 % of the genome for each mutant (Yu *et al.*, 2002). In the present work, the association of mini-Tn*5* derivative insertions and the Flp-*FRT* recombination system allowed the generation of two mutants, *P. putida* $\Delta_1$-91 and $\Delta_1$-407, deleting 240 genes which is 4.77 % of the whole genome (section 4.2.1.3). The deletion efficiency seems to be similar whether using the Cre/*loxP* system for genomic deletion in *E. coli* or the Flp-*FRT* system in *P. putida*. It is worthy to note that the bacteriophage P1 does not infect *P. putida*, thus the Cre/*loxP* system cannot be used in this strain. *E. coli* is a valued and heavily utilized organism for genetic manipulation due to the wealth of knowledge regarding its genome and physiology obtained over the last decades. The strain K12-MG1655 has a single circular chromosome of 4.64 Mbp with more than 4,700 genes. *P. putida* is also a well characterized strain, although a higher percentage of its genome remains uncharacterized. The comparison of the cellular role categories of *E. coli* K12-MG1655 and *P. putida* KT2440 reveals that around 75 % of their genes encode for six cellular roles, of which five are shared by both strains (namely ‘Hypothetical’ and ‘Conserved hypothetical proteins’, ‘Energy metabolism’, ‘Transport and binding proteins’ and ‘Cellular process’) (http://cmr.jcvi.org/cgi-bin/CMR/shared/Menu.cgi?menu=genome; JCVI CMR website). Although specific genes and/or metabolic properties have been investigated for these individual organisms (Duffy and Ford, 1997; Jeske and Altenbuchner, 2010), no information is available which compares the two. The comparison of the method combining transposon mutagenesis and site-specific recombination system between these two strains is therefore limited. It is only possible to remark the

presence of a larger deleted fragment in a single deletion step (4.1 %) in *P. putida* $\Delta_1$-407. However, this fact appeared only once during the work of the Thesis. Repetition of the experiments would elucidate the occurrence of large fragment excisions in *P. putida*. Thus, a further comparison with the results obtained in *E. coli* would be more appropriate.

Finally, another association of transposable elements with site-specific recombination system has been used in the past. In 2007, Tsuge and colleagues published their work on the construction of a library of large-segment deletion strains based on the genome of *Corynebacterium glutamicum* strain R (Tsuge *et al.*, 2007). The IS*31831* insertion sequence was used in combination with the Cre/*loxP* excision system. This approach is more similar to the method established here, with the consecutive insertion of two transposable elements carrying different exogenous DNA segments and a similar single *FRT* site. After the two transposition events and selection of the exconjugants the Cre expression vector was inserted. They were able to generate 42 independent mutants with deletion sequences ranging from 0.2 to 186 kb. By addition of all the deletions, they would have reached a reduction of 11.9 % of the total genome size.

To summarize, this method of combining transposable elements and a recombination system seems to be highly efficient for the rapid identification under established conditions of non-essential regions, as shown with the generation of the *P. putida* $\Delta_1$-407 mutant.

### 5.1.2 A method for "recycling" mini-transposons and *FRT* sites in *P. putida*

Here, mini-transposon derivatives were used in combination with the Flp-*FRT* system to create the first random genomic deletion in *P. putida*. Several independent *FRT* site combinations were randomly created in the TMT mutants as targets for the action of the Flp recombinase. *P. putida* $\Delta_1$-91 and $\Delta_1$-407 were the first two knockout mutants obtained via this method. The second part of the method was based on the development and improvement of the technique in order to delete successive genomic fragments in the same strain and generate *P. putida* $\Delta_x$ mutants. Such deletions were previously investigated in *E. coli* where independent single genomic deletions were created in different mutants by the combinatorial method previously mentioned. In the next step, the single deletions from four different mutants were cumulated in a single cell by using the transduction of phage P1, generating a final mutant with a genome reduction of 6.7 % (313.1 kb of deletion sequences) (Yu *et al.*, 2002). This transduction is the main advantage while working with *E. coli* but not with *P. putida*. For this reason, the solution was found here to simply repeat the same method as initially established. The two rounds of mini-Tn*5* derivative insertions were carried out using each deleted mutant as a new host strain and the genomic deletion was promoted by insertion of the Flp carrying vector. Even though the repetition of the procedure is efficient to generate double deleted mutants (only one month was required for the $\Delta_2$ mutant series) it remains time consuming in

comparison with the work described by Yu and colleagues. To remedy to this problem, the insertion of mini-transposons derivatives carrying an origin of transfer into the generated deleted mutants could represent a solution. This origin of transfer would allow one mutation to transfer to another deleted mutant to finally combine the deletions in a single strain, which is similar to the idea developed by Yu and colleagues.

In order to increase the yield of mutant generation and to select a final one for a specific industrial application, modifications were carried out in the second deletion step, which aimed at reducing the time and cost of the consumable. The generation of *P. putida* $\Delta_1$ SMT and TMT mutants was carried out without any intermediate step (section 4.2.2.1.2), involving that all potential candidates were automatically selected for deletion. In this way, four double-deleted mutants were generated in parallel, *P. putida* $\Delta_2$-91 col. 44 and 47 and *P. putida* $\Delta_2$-407 col. 1 and 3. This so-called "blind deletion" has the advantage of increasing the possibility of successful second deletions in one of the first mutants whilst reducing the time needed for its generation. Nonetheless, after the second deletion a supplementary step of mapping the excision points was required. The 24 possible combinations of the three *FRT* sites present in the genome of a *P. putida* $\Delta_1$ TMT mutant, before action of the Flp recombinase, shows the complexity of the situation. This is due to the absence of initial information regarding the position of the mini-Tn*5* derivatives in the $\Delta_1$ TMT candidates. Arbitrary PCR and sequencing reactions could be adapted to the specific cases; however, the exact position of the second deletion was confirmed only for one of the four double mutants obtained (*P. putida* $\Delta_2$-407 col3). The major bottleneck lies in the fact that several short sequences normally used as targets for the primers in AP-PCR are present in more exemplars in the genome due to the non-reaction of one of the mini-Tn*5* derivatives during the recombination process. To bypass this limiting point it may be possible to use an inverse PCR in order to detect the position of the second deletion as previously established for the generation of large-scale deletions (Ochman *et al.*, 1990; Goryshin *et al.*, 2003; Tsuge *et al.*, 2007). This method involves that after digestion of the genomic DNA and self-ligation of the fragments, inverse PCR is carried out that amplifies outwards from both I- and Oends. PCR products are subsequently purified and sequenced. However, due to the limited efficiency of PCRs in certain regions of the chromosome or nonspecific annealing of the primers it would be interesting also to consider the creation of a library of fragments (restricted with enzymes), clone them into a plasmid and screen them for the appropriate fragment.

The reliability of the method needs to be confirmed by the generation of mutants carrying more than two deletions in their genomes. However, the second random deletion leading to four knockout mutants was obtained within one to two months after the characterization of the $\Delta_1$ mutants, which is already a promising result for the future. Hence, the method will be applied subsequently to generate further deleted mutants and to identify the precise position of the deletions.

## 5.2 Characterizing streamlined *Pseudomonas putida* mutants

Two main, yet complementary, approaches have been employed over the last decades for the development and application of more accurate and efficient genetic tools in Eukaryotes and Prokaryotes. In the first of these approaches, the identification of genes and gene functions is carried out by single gene disruptions and analysis of the phenotype of the resulting mutants. For this purpose, genome-wide mutant libraries have been constructed via both (i) repetition of single homologous recombinations of engineered plasmids with the target genes and the resulting gene inactivation as performed in *S. cerevisiae*, *B. subtilis* and *E. coli* (Giaever *et al.*, 2002; Kobayashi *et al.*, 2003; Baba *et al.*, 2006); and (ii) establishment of mutant libraries, which were composed of mini-transposon knockout mutants, as in *P. aeruginosa* and *C. glutamicum* (Jacobs *et al.*, 2003; Suzuki *et al.*, 2006). Whilst these single-gene disruption strategies have contributed significantly to the recognition of several genes considered as essential under the given conditions of each study, they do not take into account the embedding of the single genes into the cellular regulatory networks in which they participate. The second approach for the determination of gene functions and essentiality is carried out via the generation of large-scale deletions, which allowed the excision of genomic fragments. The methods are based on homologous recombination and/or site-specific recombination systems not only in *E. coli* but also in *C. glutamicum* and *S. cerevisiae* (Kolisnychenko *et al.*, 2002; Yu *et al.*, 2002; Fukiya *et al.*, 2004; Suzuki *et al.*, 2005; Murakami *et al.*, 2007). This second approach has been used to identify mainly those genes considered as non-essential under the specified conditions if they were removed within a large fragment.

The work in this study was principally based on the second approach and has generated four double-deleted mutants with large excisions. These results are the first obtained for a *P. putida* strain. Recently, a genome-wide mutant library was published allowing a large overview of the potential essential genes under defined conditions (Molina-Henares *et al.*, 2010). Different laboratories are generating double mutants for a particular purpose aiming to study of a given gene function and protein interactions. However, the *P. putida* $\Delta_x$ mutants from the present study are the first random deleted mutants obtained via a large-scale deletion technique and with the purpose of genome streamlining.

### 5.2.1 Removal of non-persistent and conditionally non-essential genes

The establishment of SMT and TMT mutant libraries was important in two different but complementary ways. For one, all the selected mutants contributed to generate further information regarding conditionally non-essential genes. This information is valuable for the discovery of new functions for hypothetical or conserved hypothetical encoded proteins. Furthermore, the knowledge delivered by the library (mapping of the mini-Tn*5* derivatives) provides potential candidates for the

deletion of genomic fragments and subsequent information concerning conditionally non-essential genes.

***Paleome and Cenome***

Besides the notion of essentiality, the concept of persistence of a gene, indicating that the gene is present in a majority of organisms, has emerged as an important genomic feature. It is important to make the difference between both notions since a persistent gene can be non-essential and an essential gene can be non-persistent. Moreover it was shown that persistent non-essential genes and conditionally essential genes share common characteristics (Fang *et al.*, 2005). The set of persistent genes was named the *paleome*, which could be a mirror of "what was?" and "how did it happened?" at the origin of life. It was shown that the genes of the *paleome* are clustering in bacterial genomes due to strong selective pressures (Fang *et al.*, 2008). Danchin and colleagues extracted a list of around 400 persistent genes from *P. putida* strain F1 which was kindly provided for the needs of this dissertation work. Due to the high similarity between the two strains F1 and KT2440, persistent genes were identified in the genome of KT2440. Interestingly, the mini-transposons derivatives affected a persistent gene in only two cases and the three genomic deletions obtained here (*P. putida* $\Delta_1$-91, $\Delta_1$-407 and $\Delta_2$-407 col. 3) targeted the same region of the genome, which was almost not involved in the re-arrangement of the persistent genes. Even though there is a difference between essential, persistent and persistent non-essential genes (mainly related to maintenance and stress response), it seems that the cells tend to first remove parts of the *cenome* (genes encoding functions supporting life in a specific context) and *mixome* (rich in genes coding for metabolic pathways) which are not essential under defined specific conditions. This does not mean that the *paleome* represents necessarily a cold spot for transposition due to the fact that the SMT and TMT libraries were not wide enough to cover most of the genes and IR regions. The repetition of the experiments would allow a more precise assessment of the *paleome* as a hot spot.

***Mutually inclusive and exclusive mutations***

It is important to emphasize that some genes are conditionally non-essential. Previous studies have indicated that whilst some mutations can occur singly without detriment to the cell, others may become lethal only when they are combined (Yu *et al.*, 2002; Smalley *et al.*, 2003). Hence, earlier studies were able to identify "mutually inclusive" and "mutually exclusive" genes corresponding respectively to the two cases mentioned above. In 2007, Kato & Hashimoto highlighted the same ideas while removing genes from the chromosome and maintaining them in a mini-F plasmid to study the presence of essential *cis*-acting regions. They excised some particular genes, e.g. the anti-toxin genes *yefM* and *chpR*, which were considered in previous studies as being essential (Baba *et al.*, 2006). These independent excisions could occur by the simultaneous removal of two other single

genes, *yoeB* and *mazF*, respectively (Kato and Hashimoto, 2007). This demonstrates the importance of large-scale deletions for providing a better understanding of gene regulatory networks.

***Non-essentiality in P. putida***

For those cases in which the mini-transposon/s target an IR region, the flanking genes were not taken into account as the polarity of the insertion was not assessed and thus no information was available regarding the eventual disruption of the downstream gene function. Gerdes and colleagues observed in *E. coli* K12-MG1655 that insertion of a transposon upstream of an essential gene (e.g., *coaD*, *argS*) did not invoke cell death though this position in the genome was expected to be a promoter sequence or at least an influential position for the transcription of the downstream genes (Gerdes *et al.*, 2002). After analysis of the transposons, the presence of putative *E. coli*-type promoters was revealed, thus justifying the absence of a polar effect in their study. The same observations were also reported through the use of a transposon with an outward-directed promoter (Hutchison III *et al.*, 1999). Considering the absence of polarity in the present work, simple observations of the SMT and TMT mutants and their disrupted genes reveal that five out of the six predominant cellular role categories correspond to those most represented in *P. putida* KT2440, i.e. 'Regulatory functions', 'Transport and binding proteins', 'Conserved hypothetical proteins', 'Energy metabolism' and proteins of unknown function. This confirms the idea that some genes are present in several copies in the genome and thus, one of them can be removed without influencing the viability of the cell. The translated proteins can also have different substrate specificities; meaning that a protein can play different roles by recognizing diverse substrates. It can also explain that larger chromosomes are more likely to encode for a greater diversity of functions thus allowing the bacterium to adapt and respond to different environmental stress conditions (e.g. pollutants). Moreover, crucial reactions can be catalyzed by more than one isoenzyme, which, essential at the network but not at the genetic level, has been previously reported in the metabolic reconstruction of KT2440 (Puchalka *et al.*, 2008).

The genes predicted to be essential by Molina-Henares and colleagues are mainly related to amino acid, nucleotide and vitamin biosynthetic pathways. These cellular role categories correspond to the least represented (~2 %) by the disrupted genes from the SMT and TMT mutants, as expected since such deletions would be lethal for the mutants, see Figure 5-1. Some genes disrupted by either the mini-Tn*5 KpF*, the mini-Tn*5 TF* or by both were directly considered as non-essential even though no further confirmation tests were carried out. Among those genes, *lapA* (PP_0168) and *lapF* (PP_0806), which encode the two largest proteins in the genome of KT2440, were both targeted three times (see Tables 8-5 and 8-6). LapA and LapF play a role in cell adhesion and particularly in plant root colonization for the formation and maturation of the biofilm (Yousef-Coronado *et al.*, 2008; Martínez-Gil *et al.*, 2010). It is an example of conditionally non-essential genes in which cells do not require translation into these two LapA and LapF proteins for survival under laboratory conditions.

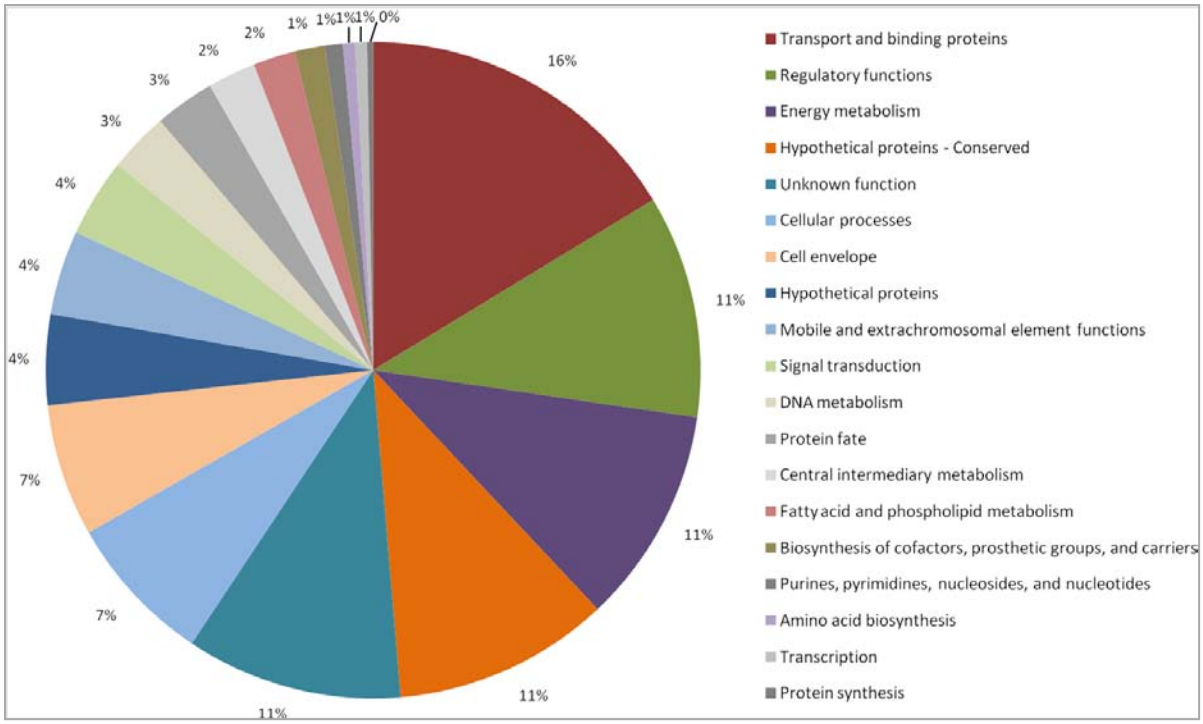



**Figure 5-1: Representation of the cellular role categories concerned by the gene disruptions in *P. putida* SMT and TMT mutants.**
The pie chart summarizes the results obtained for the SMT and TMT mutants (see Figures 4-6 and 4-8).

### 5.2.2 Highlights and implications of the genomic deletions in *P. putida* $\Delta_x$ mutants

*P. putida* $\Delta_1$-91 and $\Delta_1$-407 were obtained from *P. putida* T91 and T407, respectively. Repetition of the deletion method generated four double-deleted *P. putida* $\Delta_2$ mutants. The determination of the knockout genes was only possible for *P. putida* $\Delta_2$-407 col. 3. The main cellular categories commonly affected by all the deleted mutants (Figures 4-21 and 4-23) correspond to 'Hypothetical' and 'Conserved hypothetical proteins', 'Regulatory functions', 'Transport and binding proteins', 'Unknown functions' and 'Energy metabolism'. Five of these categories are common to those found in *P. putida* KT2440. A main part (27 %) comprises proteins whose function is not yet characterized, in concordance with the 29 % represented in the wild-type, which corresponds to the primary idea that removing functionally uncharacterized genes would avoid unexpected interference for future engineered metabolism in the selected mutant. Different characteristics and implications were revealed for each of the knockout mutants.

- **Auxotrophy**

*P. putida Δ$_x$-91 mutants*

*P. putida* $\Delta_1$-91 mutant was generated with a reduction of 0.61 % of the genome, see section 4.2.1.3.1. In total, 38 genes were removed and two genes were truncated (see Table 8-8 and Figure 4-12), which did not allow the mutant to grow in M9 medium supplemented exclusively with citrate and uracil.

Due to the involvement of gene *ilvE* in the biosynthesis and catabolism of the three branched-chain amino acids and its prediction of conditionally essential gene (Nogales *et al.*, 2008), a complementation assay was conducted for *P. putida* $\Delta_1$-91. The mutant carrying the constructed vector was able to recover the growth in M9 minimal medium supplemented only with uracil and citrate. This result is in concordance with the previous studies that show that the branched-chain amino acid aminotransferase was a constitutive enzyme responsible for the synthesis and catabolism of the three branched-chain amino acids (Marshall and Sokatch, 1972). The complementation test also confirmed that the gene *ilvE* is the only essential gene for growth of the mutant under the specified conditions, and furthermore, that no polar effect due to the mini-Tn*5* derivative insertion occurred. The influence of different amino acids on the growth of *P. putida* $\Delta_1$-91 was determined and shown in Figure 4-22, and was compared with this of the wild-type strain in M9 minimal medium supplemented only with uracil and cit. Interestingly, supplementation with L-valine and traces of L-isoleucine were not enough to recover the same growth as in the wild-type strain but sufficient to increase the fitness of the mutant. This result was unexpected due to the theoretical presence of a single and common enzyme for the synthesis and catabolism of the branched-chain amino acids. The presence of a reduced level of aminotransferase activity in an *ilvE* mutant when grown on L-valine as sole carbon source was shown (Martin *et al.*, 1973). However, this observation does not explain the growth of *P. putida* $\Delta_1$-91 in the presence of an amino acid source only. From this, it is hypothesized that another enzyme undertakes the role of the aminotransferase in its absence. The computational analysis of valine, leucine and isoleucine biosynthesis pathway pin-points the presence of a leucine dehydrogenase annotated at the same position of the pathway as the IlvE (http://www.genome.jp/kegg-bin/show_pathway?ppu00290; KEGG). Growth of the strain solely in the presence of branched-chain amino acids would indicate that the final step of transamination of 2-oxoisovalerate to yield L-valine is possible only in presence of at least one of the amino acids, mainly of L-valine. L-valine could thus be seen as an inducer of the leucine dehydrogenase (encoded by PP_4617) which is active in the synthesis and also in the catabolism of the three branched-chain amino acids.

In total, 17 deleted genes encoding for hypothetical, conserved hypothetical proteins or enzymes of unknown function were found, but none of them appeared to be essential for any metabolic or cellular process under the described conditions. Out of the 82 genes encoding for transposases five genes were deleted in *P. putida* $\Delta_1$-91. Under laboratory conditions the removal of mobile elements does not seem to affect the fitness of the bacterial cells; but rather may improve its stability.

The generation of *P. putida* $\Delta_2$-91 from $\Delta_1$-91 mutants involved also the absence of the gene *ilvE* from their genomes. However, due to the notion of mutually exclusive and inclusive mutations, it was difficult to predict if the second deletion would have a reversible effect on the auxotrophy of the

mutants. Assessment of the influence of the second deletion by comparing the growth of all mutants using the Bioscreen C MBR (Oy Growth Curves Ab Ltd, Helsinki, Finland; see section 3.6.2) did not provide sufficient evidence from which to obtain a clear conclusion regarding the growth of col. 44 and 47 in both media without the supplementation of one of the branched chain amino acids. Further experiments such as transcriptomics and proteomics under the relevant conditions (culture in M9 medium supplemented with uracil, citrate and different amino acid sources) could bring more insights into the rearrangement of the regulatory networks and possible new branches of particular metabolic pathways.

*P. putida Δx-407 mutant*

The $\Delta_x$-407 mutants were able to grow, albeit slowly in comparison with the wild-type, in M9 supplemented with citrate as a carbon source. Gene PP_3633 (*argC*) was predicted to be essential in M9 medium supplemented with glucose as the sole carbon source, in one of the *in silico* metabolic reconstructions of *P. putida* KT2440, which is related with arginine metabolism (Nogales *et al.*, 2008). Even though two *argC* genes are present in the genome of the strain (PP_0432 and PP_3633) and are thus not considered as essential for the growth of the bacteria when the function of one of them is disrupted, L-arginine was used in growth experiments. Its presence provided the evidence for the influence of the deletion on one or both ArgC enzymes, particularly since growth in liquid medium revealed that in its presence $\Delta_1$-407 mutant was able to grow faster; though slower than the wild-type (data not shown).

- **Oxidative stress**

*P. putida Δx-91 mutants*

The growth tests of the $\Delta_x$-91 mutants on medium containing hydrogen peroxide (0, 2 and 4 mM) indicated that the deletions did not seem to have affected any protein involved in oxidative stress response in comparison with *P. putida* TEC1.

*P. putida Δx-407 mutants*

Response to oxidative stress conditions was more affected in *P. putida* $\Delta_x$-407 series. The presence of the two genes PP_3587 (tpx) and PP_3668 coding for proteins involved in oxidative stress related response was already highlighted (Santos *et al.*, 2004; Kurbatov *et al.*, 2006). The removal of these genes seems to be the main reason for the behavior of the *P. putida* $\Delta_1$-407 and $\Delta_2$-407 col. 1 mutants. However, the detection of the position of the second deletion in $\Delta_2$-407 col. 1 was not successful; and thus it is difficult to draw affirmative conclusions. Complementary testing could certify this fact and determine if a single or a combination of genes influences the growth of the strain in the presence of $H_2O_2$, as with the deletion of more genes encoding proteins related with oxidative stress response (e.g. PP_2439 (ahpC) and PP_0915 (sodB)). The absence of growth of $\Delta_2$-407 col. 3 on $H_2O_2$ suggests that

the second deletion influenced even more the response to the oxidative stress. This might indicate that one of the genes encoding hypothetical, conserved hypothetical proteins or proteins of unknown function between PP_3495 and PP_3533 must have such an influence. Site-directed deletions could be applied to verify this hypothesis in the future.

- **Metabolic pathway**

Interestingly in *P. putida* $\Delta_1$-407 and subsequent double-deleted mutants, the *cat* operon (PP_3713-3716) was fully removed from their genomes. The wild-type is able to degrade benzoate by using successively the *ben* and the *cat* operon. The *ben* operon transforms benzoate into catechol which is further converted via the *ortho*-cleavage pathway, comprising the *cat* operon, to cis,cis-muconate and finally β–ketoadipate enollactone. However, absence of the cat operon would lead in theory to the inability of the strain to grow in presence of benzoate, which did not correspond to the preliminary results obtained with the Bioscreen C MBR in the presence of benzoate. The experiments would need to be repeated in order to confirm this residual growth observed for some of the mutants.

## 5.3 Towards an improvement of the general deletion procedure

The method developed in this study combined for the first time two independent known systems that were adapted for the streamlining of the *P. putida* genome. The engineered mini-transposons prepared the strain for the deletion step, which was carried out by action of the Flp site-specific recombinase. By consecutively applying rounds of transposition, target recognition and recombination, four double-deleted mutants were generated, representing the first large-scale deletions performed in a *Pseudomonas* species and a new step in the "field of random genome deletion", a field yet in its "infancy" (Suzuki *et al.*, 2008). The main conclusions and remarks of the presented work are summarized hereafter:

- The major advantage of the method is the possibility of deleting in a single step large genomic fragments, thus revealing a set of genes non-essential under specified conditions. Furthermore, no prior knowledge regarding the different gene functions is required, allowing the work with a strain in which all the gene functions have not yet been assigned.

- The establishment of intermediary SMT and TMT libraries disrupted mutants allows the user not only to reveal potential non-essential genes but also to select candidates that exhibit the best conditions for carrying out the deletion step (e.g. orientation and distance between *FRT* sites and absence of known essential genes). Although this application necessitates greater amount of time and resources for the mapping of each mini-Tn*5* derivative in the TMT mutants, the determination of the final deleted genomic fragment is automatic.

- The so-called ‘blind’ deletion method involved direct successive generation of SMT and TMT mutants subsequently treated with the Flp recombinase. Although the mapping of the deletion requires more efforts due to the lack of information about the originating TMT mutants, the procedure is obviously faster due to the absence of intermediary step. Therefore, it is worth carrying the experiments following the ‘blind’ deletion method provided that the mapping of the deletion will be modified by improving the AP-PCR method or applying inverse PCR technique.

- Two single and four double-deleted mutants were obtained as a validation of the method developed in this work. Analyses of the different mutants need to be carried out and completed in a near future. The mapping of the first deletions revealed interesting information about the behavior of the mutants when for example the branched chain amino acid aminotransferase (*P. putida* $\Delta_1$-91) or the full *cat* operon (*P. putida* $\Delta_1$-407) is removed. In the next steps analyses such as transcriptomics and proteomics could be used in order to extract data from the mutants and bring new insights into the wild-type strain.

Despite of these interesting results, improvements of the different steps can be made in order to generate a high-throughput random deletion system, which could be applied to several *Pseudomonas* species. In a first step, the same protocol could be directly applied to *P. putida* KT2440 by removing only the *pyrF* operon from mini-Tn*5 KpF*. It is believed that in this case the efficiency of the final site-specific recombination would not be changed. Improvement of the method targets different modules: (i) increase of the number of different markers by using other antibiotics or recognition cassettes (e.g. *lacZ, gfp*), which avoids several insertions of the same mini-Tn*5* or (ii) transformation with linear mini-Tn*5* derivatives, which allows them to be inserted in a wider range of hosts. Furthermore, a combination of the present established method with site-directed mutagenesis would allow the user to manipulate a reduced mutant to target a specific gene. For this purpose, a set of engineered suicide vectors was designed in combination with the Flp-*FRT* system also and need to be improved in a near future for the application on the obtained double mutants. Finally, the fitness of the strain is modified through the different deletions; however, it was already shown for *E. coli* that by removing some genes the mutant has the tendency to grow better (Hashimoto *et al.*, 2005), which might happen for *P. putida* if particular genes are removed from the genome.

Generation of the double mutants opens the way for further reduction of the genome and through the removal of non-essential and unstable elements provides a future of re-engineered strains of greater industrial use. It is possible to take into account a potential automation of the deletion procedure in order (i) to reach in a faster way a further reduced genome and thus, to be able to endow with the new genetic circuits and (ii) to offer a wider panel of streamlined mutants adapted to the production of different high valuable compounds for biotechnological applications such as the biotransformation of

aromatic compounds into high-added value products or production of precursors of bioplastics from glycerol.

# 6. Summary

*Pseudomonas putida* KT2440 is a certified biosafety strain able of degrading aromatic compounds. Moreover, it is an ideal host for heterologous gene expression which allows the expansion of the range of chemicals which can be degraded or further converted into high value compounds. These features were the main reasons for the choice of *P. putida* TEC1 (derived strain from KT2440) as candidate strain for the streamlining of its genome under defined conditions. The present work allowed the elimination of unecessary genomic traits, possibly enabling a simplification of the metabolic and regulatory hierarchies underlying cellular processes. This will ultimately lead to a better understanding and control of the strain.

Transposon mutagenesis is a common tool used for the generation of knockout mutants aiming at the determination of gene functions or the improvement of specific metabolic or cellular pathways. Here, transposon mutagenesis was used to randomly remove DNA fragments allowing faster reduction of the number of genes with less manipulations than through the repeated site-directed mutagenesis method. In the first part of the work, a set of two modified mini-Tn*5* transposons was randomly inserted in the chromosome of *P. putida* and two libraries of disrupted mutants were built. In the second part, the *Saccharomyces cerevisiae* Flp-*FRT* recombination system was associated with transposon mutagenesis in order to proceed with the final step of random excision of genomic fragments in *P. putida* TEC1. The Flp-*FRT* system is a powerfull tool for recombination between two specific sites and subsequent removal of the fragment situated in-between. This tool was previously used in different studies but was here applied for the first time to a *P. putida* strain for large-scale deletions.

The streamlining process involved the reduction of the 6.18 Mb genome through a combination of random-transposon mutagenesis and site-specific recombinations, which led to large-scale deletions. In total six mutants were obtained: (i) two $\Delta_1$ mutants in which an independent genomic fragment was removed (0.6 and 4.1 % of genome size reduction) and (ii) four $\Delta_2$ mutants (4.55 % of reduction for *P. putida* $\Delta_2$-407 col.3, with the mapped deletion). From the mapped deleted mutants, 244 genes were removed in total.

DNA fragments non-essential for the survival of the bacteria under the given laboratory conditions were removed within the two single and four double mutants obtained via the application of the deletion system. Genetic characterizations allowed the determination of non-essential fragments in both single-deleted strains. Even though a important part comprised genes whose function is not yet characterized, some excised genes are involved in amino acid biosynthesis, benzoate degradation or oxidative stress response. Subsequent phenotypic tests identified auxotrophy for some of the mutants

under the specified conditions. Auxotrophy on branched-chain amino acids and the degradation of benzoate in absence of the cat operon are main points of interests for a future axe of research of the produced mutants.

The resulting exicisions aim at generating a streamlined version of the *P. putida* genome, thereby paving the way for a better understanding of the metabolic and regulatory wiring of the cells, and the subsequent re-engineering of this bacterium for biocatalysis.

# 7. References

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-402.

Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner and H. Mori (2006). "Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection." Mol Syst Biol **2**: 2006 0008.

Bayley, S. A., C. J. Duggleby, M. J. Worsey, P. A. Williams, K. G. Hardy and P. Broda (1977). "Two modes of loss of the Tol function from Pseudomonas putida mt-2." Mol Gen Genet **154**(2): 203-4.

Benner, S. A. and A. M. Sismour (2005). "Synthetic biology." Nat Rev Genet **6**(7): 533-43.

Berg, D. E. (1989). Transposon Tn5. Mobile DNA. D. E. B. a. M. M. H. (ed.). Washington, D.C., American Society for Microbiology**:** p. 183-208.

Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau and Y. Shao (1997). "The complete genome sequence of Escherichia coli K-12." Science **277**(5331): 1453-62.

Bodey, G. P., R. Bolivar, V. Fainstein and L. Jadeja (1983). "Infections caused by Pseudomonas aeruginosa." Rev Infect Dis **5**(2): 279-313.

Cases, I. and V. de Lorenzo (1998). "Expression systems and physiological control of promoter activity in bacteria." Curr Opin Microbiol **1**(3): 303-10.

Chakrabarty, A. M. (1976). "Plasmids in Pseudomonas." Annu Rev Genet **10**: 7-30.

Cherepanov, P. P. and W. Wackernagel (1995). "Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant." Gene **158**(1): 9-14.

Chiarabelli, C., P. Stano and P. L. Luisi (2009). "Chemical approaches to synthetic biology." Curr Opin Biotechnol **20**(4): 492-7.

Clarke, P. H. (1982). "The metabolic versatility of pseudomonads." Antonie Van Leeuwenhoek **48**(2): 105-30.

Courvalin, P. and C. Carlier (1986). "Transposable multiple antibiotic resistance in Streptococcus pneumoniae." Mol Gen Genet **205**(2): 291-7.

Craig, N. L. (1996). "Transposon Tn7." Curr Top Microbiol Immunol **204**: 27-48.

# REFERENCES


Dam, P., V. Olman, K. Harris, Z. Su and Y. Xu (2007). "Operon prediction using both genome-specific and general genomic information." Nucleic Acids Res **35**(1): 288-98.

Datsenko, K. A. and B. L. Wanner (2000). "One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products." Proc Natl Acad Sci U S A **97**(12): 6640-5.

Davies, C. J. and C. A. Hutchison, 3rd (1995). "Insertion site specificity of the transposon Tn3." Nucleic Acids Res **23**(3): 507-14.

De las Heras, A., C. A. Carreno and V. De Lorenzo (2008). "Stable implantation of orthogonal sensor circuits in Gram-negative bacteria for environmental release." Environ Microbiol **10**(12): 3305-16.

De Lorenzo, V., L. Eltis, B. Kessler and K. N. Timmis (1993). "Analysis of Pseudomonas gene products using lacI(q)/Ptrp-lac plasmids and transposons that confer conditional phenotypes." Gene **123**(1): 17-24.

De Lorenzo, V., M. Herrero, U. Jakubzik and K. N. Timmis (1990). "Mini-Tn5 transposoon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria." Journal of Bacteriology **172**(11): 6568-6572.

de Lorenzo, V., L. Serrano and A. Valencia (2006). "Synthetic biology: challenges ahead." Bioinformatics **22**(2): 127-8.

De Lorenzo, V. and K. N. Timmis (1994). "Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons." Methods in Enzymology **235**: 386-405.

Duffy, K. J. and R. M. Ford (1997). "Turn angle and run time distributions characterize swimming behavior for Pseudomonas putida." J Bacteriol **179**(4): 1428-30.

Duque, E., A. J. Molina-Henares, J. De la Torre, M. A. Molina-Henares, T. Del Castillo, J. Lam and J. L. Ramos (2007). Towards a genome-wide mutant library of Pseudomonas putida strain KT2440. Pseudomonas - A Model System in Biology. J. L. Ramos and A. Filloux. **5:** 227-251.

Fang, G., E. Rocha and A. Danchin (2005). "How essential are nonessential genes?" Mol Biol Evol **22**(11): 2147-56.

Fang, G., E. P. Rocha and A. Danchin (2008). "Persistence drives gene clustering in bacterial genomes." BMC Genomics **9**: 4.

Forster, A. C. and G. M. Church (2006). "Towards synthesis of a minimal cell." Mol Syst Biol **2**: 45.

Franklin, F. C., M. Bagdasarian, M. M. Bagdasarian and K. N. Timmis (1981). "Molecular and functional analysis of the TOL plasmid pWWO from Pseudomonas putida and cloning of genes for the entire regulated aromatic ring meta cleavage pathway." Proc Natl Acad Sci U S A **78**(12): 7458-62.

Fukiya, S., H. Mizoguchi and H. Mori (2004). "An improved method for deleting large regions of Escherichia coli K-12 chromosome using a combination of Cre/loxP and lambda Red." FEMS Microbiol Lett **234**(2): 325-31.

Galvao, T. C. and V. de Lorenzo (2005). "Adaptation of the yeast URA3 selection system to gram-negative bacteria and generation of a {delta}betCDE Pseudomonas putida strain." Appl Environ Microbiol **71**(2): 883-92.

Gawron-Burke, C. and D. B. Clewell (1982). "A transposon in Streptococcus faecalis with fertility properties." Nature **300**(5889): 281-4.

Gerdes, S. Y., M. D. Scholle, M. D'Souza, A. Bernal, M. V. Baev, M. Farrell, O. V. Kurnasov, M. D. Daugherty, F. Mseeh, B. M. Polanuyer, J. W. Campbell, S. Anantha, K. Y. Shatalin, S. A. Chowdhury, M. Y. Fonstein and A. L. Osterman (2002). "From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways." J Bacteriol **184**(16): 4555-72.

Giaever, G., A. M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila, K. Anderson, B. Andre, A. P. Arkin, A. Astromoff, M. El-Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K. D. Entian, P. Flaherty, F. Foury, D. J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J. H. Hegemann, S. Hempel, Z. Herman, D. F. Jaramillo, D. E. Kelly, S. L. Kelly, P. Kotter, D. LaBonte, D. C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S. L. Ooi, J. L. Revuelta, C. J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D. D. Shoemaker, S. Sookhai-Mahadeo, R. K. Storms, J. N. Strathern, G. Valle, M. Voet, G. Volckaert, C. Y. Wang, T. R. Ward, J. Wilhelmy, E. A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J. D. Boeke, M. Snyder, P. Philippsen, R. W. Davis and M. Johnston (2002). "Functional profiling of the Saccharomyces cerevisiae genome." Nature **418**(6896): 387-91.

Gibson, D. G., J. I. Glass, C. Lartigue, V. N. Noskov, R. Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z. Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison, 3rd, H. O. Smith and J. C. Venter "Creation of a bacterial cell controlled by a chemically synthesized genome." Science **329**(5987): 52-6.

Gidoni, D., M. Bar and N. Gilboa (2001). "FLP/FRT-mediated restoration of normal phenotypes and clonal sectors formation in rolC transgenic tobacco." Transgenic Res **10**(4): 317-28.

Gilbert, E. S., A. W. Walker and J. D. Keasling (2003). "A constructed microbial consortium for biodegradation of the organophosphorus insecticide parathion." Appl Microbiol Biotechnol **61**(1): 77-81.

Goryshin, I. Y., T. A. Naumann, J. Apodaca and W. S. Reznikoff (2003). "Chromosomal deletion formation system based on Tn5 double transposition: use for making minimal genomes and essential gene analysis." Genome Res **13**(4): 644-53.

Goryshin, I. Y. and W. S. Reznikoff (1998). "Tn5 in vitro transposition." J Biol Chem **273**(13): 7367-74.

Goss, W. A., W. H. Deitz and T. M. Cook (1965). "Mechanism of Action of Nalidixic Acid on Escherichia Coli.Ii. Inhibition of Deoxyribonucleic Acid Synthesis." J Bacteriol **89**: 1068-74.

Hashimoto, M., T. Ichimura, H. Mizoguchi, K. Tanaka, K. Fujimitsu, K. Keyamura, T. Ote, T. Yamakawa, Y. Yamazaki, H. Mori, T. Katayama and J. Kato (2005). "Cell size and nucleoid organization of engineered Escherichia coli cells with a reduced genome." Mol Microbiol **55**(1): 137-49.

Herrero, M., V. De Lorenzo and K. N. Timmis (1990). "Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria." Journal of Bacteriology **172**(11): 6557-6567.

Hoang, T. T., R. R. Karkhoff-Schweizer, A. J. Kutchma and H. P. Schweizer (1998). "A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa mutants." Gene **212**(1): 77-86.

Hoang, T. T., A. J. Kutchma, A. Becher and H. P. Schweizer (2000). "Integration-proficient plasmids for Pseudomonas aeruginosa: site-specific integration and use for engineering of reporter and expression strains." Plasmid **43**(1): 59-72.

Holtwick, R., A. von Wallbrunn, H. Keweloh and F. Meinhardt (2001). "A novel rolling-circle-replicating plasmid from Pseudomonas putida P8: molecular characterization and use as vector." Microbiology **147**(Pt 2): 337-44.

Huang, L. C., E. A. Wood and M. M. Cox (1991). "A bacterial model system for chromosomal targeting." Nucleic Acids Res **19**(3): 443-8.

Huang, L. C., E. A. Wood and M. M. Cox (1997). "Convenient and reversible site-specific targeting of exogenous DNA into a bacterial chromosome by use of the FLP recombinase: the FLIRT system." J Bacteriol **179**(19): 6076-83.

Husson, R. N., B. E. James and R. A. Young (1990). "Gene replacement and expression of foreign DNA in mycobacteria." J Bacteriol **172**(2): 519-24.

Hutchison III, C. A., S. N. Peterson, S. R. Gill, R. T. Cline, O. White, C. M. Fraser, H. O. Smith and J. C. Venter (1999). "Global transposon mutagenesis and a minimal mycoplasma genome." Science **286**(5447): 2165-2169.

Itaya, M. (1995). "An estimation of minimal genome size required for life." FEBS Lett **362**(3): 257-60.

Jacobs, M. A., A. Alwood, I. Thaipisuttikul, D. Spencer, E. Haugen, S. Ernst, O. Will, R. Kaul, C. Raymond, R. Levy, L. Chun-Rong, D. Guenthner, D. Bovee, M. V. Olson and C. Manoil (2003). "Comprehensive transposon mutant library of Pseudomonas aeruginosa." Proc Natl Acad Sci U S A **100**(24): 14339-44.

Jayaraman, M., I. Grainge and G. Tribble (2002). Site-Specific Recombination by the Flp Protein of *Saccharomyces cerevisiae*. Mobile DNA II. N. Craig, R. Craigie, M. Gellert and A. M. Lambowitz, ASM Press**:** p.192-218.

Jeske, M. and J. Altenbuchner (2010). "The Escherichia coli rhamnose promoter rhaP(BAD) is in Pseudomonas putida KT2440 independent of Crp-cAMP activation." Appl Microbiol Biotechnol **85**(6): 1923-33.

Jimenez, J. I., B. Minambres, J. L. Garcia and E. Diaz (2002). "Genomic analysis of the aromatic catabolic pathways from Pseudomonas putida KT2440." Environ Microbiol **4**(12): 824-41.

Jimenez, J. I., B. Minambres, J. L. Garcia and E. Díaz (2004). Genomic insights in the metabolism of aromatic compounds in *Pseudomonas*. Pseudomonas, Vol III. J. L. Ramos. New York, Kluwer Academic/Plenum Publishers**:** p. 425-462.

Johnson, R. C., J. C. Yin and W. S. Reznikoff (1982). "Control of Tn5 transposition in Escherichia coli is mediated by protein from the right repeat." Cell **30**(3): 873-82.

Kanehisa, M., S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki and M. Hirakawa (2006). "From genomics to chemical genomics: new developments in KEGG." Nucleic Acids Res **34**(Database issue): D354-7.

Kato, J. and M. Hashimoto (2007). "Construction of consecutive deletions of the Escherichia coli chromosome." Mol Syst Biol **3**: 132.

Knipfer, N., A. Seth and T. E. Shrader (1997). "Unmarked gene integration into the chromosome of Mycobacterium smegmatis via precise replacement of the pyrF gene." Plasmid **37**(2): 129-40.

Kobayashi, K., S. D. Ehrlich, A. Albertini, G. Amati, K. K. Andersen, M. Arnaud, K. Asai, S. Ashikaga, S. Aymerich, P. Bessieres, F. Boland, S. C. Brignell, S. Bron, K. Bunai, J. Chapuis, L. C. Christiansen, A. Danchin, M. Debarbouille, E. Dervyn, E. Deuerling, K. Devine, S. K. Devine, O. Dreesen, J. Errington, S. Fillinger, S. J. Foster, Y. Fujita, A. Galizzi, R. Gardan, C. Eschevins, T. Fukushima, K. Haga, C. R. Harwood, M. Hecker, D. Hosoya, M. F. Hullo, H. Kakeshita, D. Karamata, Y. Kasahara, F. Kawamura, K. Koga, P. Koski, R. Kuwana, D. Imamura, M. Ishimaru, S. Ishikawa, I. Ishio, D. Le Coq, A. Masson, C. Mauel, R. Meima, R. P. Mellado, A. Moir, S. Moriya, E. Nagakawa, H. Nanamiya, S. Nakai, P. Nygaard, M. Ogura, T. Ohanan, M. O'Reilly, M. O'Rourke, Z. Pragai, H. M. Pooley, G. Rapoport, J. P. Rawlins, L. A. Rivas, C. Rivolta, A. Sadaie, Y. Sadaie, M. Sarvas, T. Sato, H. H. Saxild, E. Scanlan, W. Schumann, J. F. Seegers, J. Sekiguchi, A. Sekowska, S. J. Seror, M. Simon, P. Stragier, R. Studer, H. Takamatsu, T. Tanaka, M. Takeuchi, H. B. Thomaides, V. Vagner, J. M. van Dijl, K. Watabe, A. Wipat, H. Yamamoto, M. Yamamoto, Y. Yamamoto, K. Yamane, K. Yata, K. Yoshida, H. Yoshikawa, U. Zuber and N. Ogasawara (2003). "Essential Bacillus subtilis genes." Proc Natl Acad Sci U S A **100**(8): 4678-83.

Kojima, Y., H. Fujisawa, A. Nakazawa, T. Nakazawa, F. Kanetsuna, H. Taniuchi, M. Nozaki and O. Hayaishi (1967). "Studies on pyrocatechase. I. Purification and spectral properties." J Biol Chem **242**(14): 3270-8.

Kolisnychenko, V., G. Plunkett, 3rd, C. D. Herring, T. Feher, J. Posfai, F. R. Blattner and G. Posfai (2002). "Engineering a reduced Escherichia coli genome." Genome Res **12**(4): 640-7.

Kolter, R. (1981). "Replication properties of plasmic R6K." Plasmid **5**(1): 2-9.

Koonin, E. V. (2000). "How many genes can make a cell: the minimal-gene-set concept." Annu Rev Genomics Hum Genet **1**: 99-116.

Kopke, K., B. Hoff and U. Kuck (2010). "Application of the yeast FLP/FRT recombinase system in filamentous fungi: Marker recycling and construction of knockout strains devoid of heterologous genes." Appl Environ Microbiol.

Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop, 2nd and K. M. Peterson (1995). "Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes." Gene **166**(1): 175-6.

Kurbatov, L., D. Albrecht, H. Herrmann and L. Petruschka (2006). "Analysis of the proteome of Pseudomonas putida KT2440 grown on different sources of carbon and energy." Environ Microbiol **8**(3): 466-78.

Leduc, S. (1912). La biologie synthétique, étude de biophysique. Paris.

Lehoux, D. E., F. Sanschagrin and R. C. Levesque (2002). "Identification of in vivo essential genes from Pseudomonas aeruginosa by PCR-based signature-tagged mutagenesis." FEMS Microbiology Letters **210**(1): 73-80.

Liberati, N. T., J. M. Urbach, S. Miyata, D. G. Lee, E. Drenkard, G. Wu, J. Villanueva, T. Wei and F. M. Ausubel (2006). "An ordered, nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants." Proc Natl Acad Sci U S A **103**(8): 2833-8.

Mahillon, J. and M. Chandler (1998). "Insertion sequences." Microbiol Mol Biol Rev **62**(3): 725-74.

Mao, F., P. Dam, J. Chou, V. Olman and Y. Xu (2009). "DOOR: a database for prokaryotic operons." Nucleic Acids Res **37**(Database issue): D459-63.

Marshall, V. D. and J. R. Sokatch (1972). "Regulation of valine catabolism in Pseudomonas putida." Journal of Bacteriology **110**(3): 1073-1081.

Martin, R. R., V. D. Marshall, J. R. Sokatch and L. Unger (1973). "Common enzymes of branched chain amino acid catabolism in Pseudomonas putida." Journal of Bacteriology **115**(1): 198-204.

Martínez-Gil, M., F. Yousef-Coronado and M. Espinosa-Urgel (2010). "LapF, the second largest Pseudomonas putida protein, contributes to plant root colonization and determines biofilm architecture." Molecular Microbiology **77**(3): 549-561.

Meselson, M. and R. Yuan (1968). "DNA restriction enzyme from E. coli." Nature **217**(5134): 1110-4.

Molina-Henares, M. A., J. de la Torre, A. Garcia-Salamanca, A. J. Molina-Henares, M. C. Herrera, J. L. Ramos and E. Duque (2010). "Identification of conditionally essential genes for growth of Pseudomonas putida KT2440 on minimal medium through the screening of a genome-wide mutant library." Environ Microbiol.

Murakami, K., E. Tao, Y. Ito, M. Sugiyama, Y. Kaneko, S. Harashima, T. Sumiya, A. Nakamura and M. Nishizawa (2007). "Large scale deletions in the Saccharomyces cerevisiae genome create strains with altered regulation of carbon metabolism." Appl Microbiol Biotechnol **75**(3): 589-97.

Mushegian, A. R. and E. V. Koonin (1996). "A minimal gene set for cellular life derived by comparison of complete bacterial genomes." Proc Natl Acad Sci U S A **93**(19): 10268-73.

Nakayama, K., S. M. Kelly and R. Curtiss Iii (1988). "Construction of an ASD<sup>+</sup> expression-cloning vector: Stable maintenance and high level expression of cloned genes in a Salmonella vaccine strain." Bio/Technology **6**(6): 693-697.

Nakazawa, T. (2002). "Travels of a Pseudomonas, from Japan around the world." Environ Microbiol **4**(12): 782-6.

Nelson, K. E., C. Weinel, I. T. Paulsen, R. J. Dodson, H. Hilbert, V. A. Martins dos Santos, D. E. Fouts, S. R. Gill, M. Pop, M. Holmes, L. Brinkac, M. Beanan, R. T. DeBoy, S. Daugherty, J. Kolonay, R. Madupu, W. Nelson, O. White, J. Peterson, H. Khouri, I. Hance, P. Chris Lee, E. Holtzapple, D. Scanlan, K. Tran, A. Moazzez, T. Utterback, M. Rizzo, K. Lee, D. Kosack, D. Moestl, H. Wedler, J. Lauber, D. Stjepandic, J. Hoheisel, M. Straetz, S. Heim, C. Kiewitz, J. A. Eisen, K. N. Timmis, A. Dusterhoft, B. Tummler and C. M. Fraser (2002). "Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440." Environ Microbiol **4**(12): 799-808.

Nogales, J., B. O. Palsson and I. Thiele (2008). "A genome-scale metabolic reconstruction of Pseudomonas putida KT2440: iJN746 as a cell factory." BMC Syst Biol **2**: 79.

O'Toole, G. A., L. A. Pratt, P. I. Watnick, D. K. Newman, V. B. Weaver and R. Kolter (1999). "Genetic approaches to study of biofilms." Methods Enzymol **310**: 91-109.

Ochman, H., J. W. Ajioka, D. Garza and D. L. Hartl (1990). "Inverse polymerase chain reaction." Biotechnology (N Y) **8**(8): 759-60.

Olivera, E. R., D. Carnicero, R. Jodra, B. Minambres, B. Garcia, G. A. Abraham, A. Gallardo, J. S. Roman, J. L. Garcia, G. Naharro and J. M. Luengo (2001). "Genetically engineered Pseudomonas: a factory of new bioplastics with broad applications." Environ Microbiol **3**(10): 612-8.

Papin, J. A., N. D. Price, S. J. Wiback, D. A. Fell and B. O. Palsson (2003). "Metabolic pathways in the post-genome era." Trends Biochem Sci **28**(5): 250-8.

Posfai, G., M. D. Koob, H. A. Kirkpatrick and F. R. Blattner (1997). "Versatile insertion plasmids for targeted genome manipulations in bacteria: isolation, deletion, and rescue of the pathogenicity island LEE of the Escherichia coli O157:H7 genome." J Bacteriol **179**(13): 4426-8.

Pósfai, G., G. Plunkett Iii, T. Fehér, D. Frisch, G. M. Keil, K. Umenhoffer, V. Kolisnychenko, B. Stahl, S. S. Sharma, M. De Arruda, V. Burland, S. W. Harcum and F. R. Blattner (2006). "Emergent properties of reduced-genome Escherichia coli." Science **312**(5776): 1044-1046.

Puchalka, J., M. A. Oberhardt, M. Godinho, A. Bielecka, D. Regenhardt, K. N. Timmis, J. A. Papin and V. A. Martins dos Santos (2008). "Genome-scale reconstruction and analysis of the Pseudomonas putida KT2440 metabolic network facilitates applications in biotechnology." PLoS Comput Biol **4**(10): e1000210.

Puttamreddy, S., N. A. Cornick and F. C. Minion (2010). "Genome-wide transposon mutagenesis reveals a role for pO157 genes in biofilm development in Escherichia coli O157:H7 EDL933." Infect Immun **78**(6): 2377-84.

Ramos, J. L., S. Marques and K. N. Timmis (1997). "Transcriptional control of the Pseudomonas TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators." Annu Rev Microbiol **51**: 341-73.

Ramos, J. L., A. Wasserfallen, K. Rose and K. N. Timmis (1987). "Redesigning metabolic routes: manipulation of TOL plasmid pathway for catabolism of alkylbenzoates." Science **235**(4788): 593-6.

Rasmussen, S., M. A. Bedau, L. Chen, D. Deamer, D. C. Krakauer, N. H. Packard and P. F. Stadler (2008). Protocells: Bridging Nonliving and Living Matter, MIT Press.

Reineke, W. (1998). "Development of hybrid strains for the mineralization of chloroaromatics by patchwork assembly." Annu Rev Microbiol **52**: 287-331.

Reznikoff, W. S. (2002). Tn5 Transposition. Mobile DNA II. N. L. Craig, R. Craigie, M. Gellert and A. M. Lambowitz. Washington, ASM Press**:** p. 403-422.

Reznikoff, W. S. (2003). "Tn5 as a model for understanding DNA transposition." Mol Microbiol **47**(5): 1199-206.

Ross-Macdonald, P., P. S. R. Coelho, T. Roemer, S. Agarwal, A. Kumar, R. Jansen, K. H. Cheung, A. Sheehan, D. Symonlatis, L. Umansky, M. Heldtman, F. K. Nelson, H. Iwasaki, K. Hager, M. Gerstein, P. Miller, G. S. Roeder and M. Snyder (1999). "Large-scale analysis of the yeast genome by transposon tagging and gene disruption." Nature **402**(6760): 413-418.

Rozen, S. and H. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers." Methods Mol Biol **132**: 365-86.

Sadowski, P. D. (1995). "The Flp recombinase of the 2-microns plasmid of Saccharomyces cerevisiae." Prog Nucleic Acid Res Mol Biol **51**: 53-91.

Sambrook, J., T. Maniatis and T. Fritsch (1989). Molecular cloning: a laboratory manual. N. Y., Cold Spring Harbor Laboratory, Cold Spring Harbor.

Sanchez-Romero, J. M., R. Diaz-Orejas and V. De Lorenzo (1998). "Resistance to tellurite as a selection marker for genetic manipulations of Pseudomonas strains." Appl Environ Microbiol **64**(10): 4040-6.

Santos, P. M., D. Benndorf and I. Sa-Correia (2004). "Insights into Pseudomonas putida KT2440 response to phenol-induced stress by quantitative proteomics." Proteomics **4**(9): 2640-52.

Sauer, B. (1987). "Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae." Mol Cell Biol **7**(6): 2087-96.

Schmid, A., J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts and B. Witholt (2001). "Industrial biocatalysis today and tomorrow." Nature **409**(6817): 258-68.

Schweizer, H. P. (2003). "Applications of the Saccharomyces cerevisiae Flp-FRT system in bacterial genetics." J Mol Microbiol Biotechnol **5**(2): 67-77.

Shevchenko, Y., G. G. Bouffard, Y. S. Butterfield, R. W. Blakesley, J. L. Hartley, A. C. Young, M. A. Marra, S. J. Jones, J. W. Touchman and E. D. Green (2002). "Systematic sequencing of cDNA clones using the transposon Tn5." Nucleic Acids Res **30**(11): 2469-77.

Smalley, D. J., M. Whiteley and T. Conway (2003). "In search of the minimal Escherichia coli genome." Trends Microbiol **11**(1): 6-8.

Sole, R. V., A. Munteanu, C. Rodriguez-Caso and J. Macia (2007). "Synthetic protocell biology: from reproduction to computation." Philos Trans R Soc Lond B Biol Sci **362**(1486): 1727-39.

Stachel, S. E., G. An, C. Flores and E. W. Nester (1985). "A Tn3 lacZ transposon for the random generation of beta-galactosidase gene fusions: application to the analysis of gene expression in Agrobacterium." EMBO J **4**(4): 891-8.

Stein, D. C. (1992). "Plasmids with easily excisable xylE cassettes." Gene **117**(1): 157-8.

Storici, F., M. Coglievina and C. V. Bruschi (1999). "A 2-microm DNA-based marker recycling system for multiple gene disruption in the yeast Saccharomyces cerevisiae." Yeast **15**(4): 271-83.

Stothard, P., G. Van Domselaar, S. Shrivastava, A. Guo, B. O'Neill, J. Cruz, M. Ellison and D. S. Wishart (2005). "BacMap: an interactive picture atlas of annotated bacterial genomes." Nucleic Acids Res **33**(Database issue): D317-20.

Suzuki, N., M. Inui and H. Yukawa (2008). "Random genome deletion methods applicable to prokaryotes." Appl Microbiol Biotechnol **79**(4): 519-26.

Suzuki, N., N. Okai, H. Nonaka, Y. Tsuge, M. Inui and H. Yukawa (2006). "High-throughput transposon mutagenesis of Corynebacterium glutamicum and construction of a single-gene disruptant mutant library." Appl Environ Microbiol **72**(5): 3750-5.

Suzuki, N., S. Okayama, H. Nonaka, Y. Tsuge, M. Inui and H. Yukawa (2005). "Large-scale engineering of the Corynebacterium glutamicum genome." Appl Environ Microbiol **71**(6): 3369-72.

Szybalski, W. (1974). "In vivo and in vitro initiation of transcription." Adv Exp Med Biol **44**(1): 23-4.

Thomas, C. M. and C. A. Smith (1987). "Incompatibility group P plasmids: genetics, evolution, and use in genetic manipulation." Annu Rev Microbiol **41**: 77-101.

Timmis, K., F. Cabello and S. N. Cohen (1975). "Cloning, isolation, and characterization of replication regions of complex plasmid genomes." Proc Natl Acad Sci U S A **72**(6): 2242-6.

Timmis, K. N., R. J. Steffan and R. Unterman (1994). "Designing microorganisms for the treatment of toxic wastes." Annu Rev Microbiol **48**: 525-57.

Toyota, T., N. Maru, M. M. Hanczyc, T. Ikegami and T. Sugawara (2009). "Self-propelled oil droplets consuming "fuel" surfactant." J Am Chem Soc **131**(14): 5012-3.

Tricot, C., V. Stalon and C. Legrain (1991). "Isolation and characterization of Pseudomonas putida mutants affected in arginine, ornithine and citrulline catabolism: function of the arginine oxidase and arginine succinyltransferase pathways." J Gen Microbiol **137**(12): 2911-8.

Tsuge, Y., N. Suzuki, M. Inui and H. Yukawa (2007). "Random segment deletion based on IS31831 and Cre/loxP excision system in Corynebacterium glutamicum." Appl Microbiol Biotechnol **74**(6): 1333-41.

Wackett, L. P. (2003). "Pseudomonas putida--a versatile biocatalyst." Nat Biotechnol **21**(2): 136-8.

Weinel, C., K. E. Nelson and B. Tummler (2002). "Global features of the Pseudomonas putida KT2440 genome sequence." Environ Microbiol **4**(12): 809-18.

Williams, P. A. and K. Murray (1974). "Metabolism of benzoate and the methylbenzoates by Pseudomonas putida (arvilla) mt-2: evidence for the existence of a TOL plasmid." J Bacteriol **120**(1): 416-23.

Wubbolts, M. G. and K. N. Timmis (1990). "Biotransformation of substituted benzoates to the corresponding cis-diols by an engineered strain of Pseudomonas oleovorans producing the TOL plasmid-specified enzyme toluate-1,2-dioxygenase." Appl Environ Microbiol **56**(2): 569-71.

Yousef-Coronado, F., M. L. Travieso and M. Espinosa-Urgel (2008). "Different, overlapping mechanisms for colonization of abiotic and plant surfaces by Pseudomonas putida." FEMS Microbiol Lett **288**(1): 118-24.

Yu, B. J., B. H. Sung, M. D. Koob, C. H. Lee, J. H. Lee, W. S. Lee, M. S. Kim and S. C. Kim (2002). "Minimization of the Escherichia coli genome using a Tn5-targeted Cre/loxP excision system." Nat Biotechnol **20**(10): 1018-23.

# 8. Appendix

## 8.1 List of primers

**Table 8-1: List of primers used for the construction of the mini-Tn*5* derivatives**

| Primers | Sequence (5'-3') | Melting Temperature (°C) | Reference |
|---|---|---|---|
| *pyrF* 1F | cg*GGATCC*GTCGGGCCGTGGGCAACG | 74.5 | This work |
| *pyrF* 2R | cccgAAGCTTTTACCCACGGATCTCCGC | 71.0 | This work |
| Fp 1F(*Not*I) | ataagaat*GCGGCCGC*GTTGACAAAGGGAATCAG GG | 72.65 | This work |
| Fp 2R | ataagaatGCGGCCGCAAGCTTTTAC | 64.59 | This work |
| F 10F | ATTCCGGGCATTGCTGTTGAC | 59.8 | This work |
| F 11R | GGCCTGGCACCTATAATTGAACC | 64.55 | This work |

**Table 8-2: List of primers used for the AP-PCR**

| Primers | Sequence (5'-3') | Melting Temperature (°C) | Reference |
|---|---|---|---|
| ARB1 | GGCCACGCGTCGACTAGTACNNNNNNNNNN GATAT | 73.0 | (O'Toole *et al.*, 1999) |
| ARB2 | GGCCACGCGTCGACTAGTAC | 63.5 | (O'Toole *et al.*, 1999) |
| APPCR-Km1 | CTACAGCTCGTTTCACGCTGAATA | 61.0 | This work |
| APPCR-Km3 | CTTGTGCAATGTAACATCAGAG | 56.5 | This work |
| kilA rev | ACGCTTTGTTCTTCCATTCG | 55.3 | This work |
| Iend Rev2 | CTGTCTCTTATACACATC | 55.9 | This work |
| Oend F | ACTTGTGTATAAGAGTCAGT | 51.2 | This work |
| F 8R(*NotI*) | TAAGAAT*GCGGCCGC*GAAGTTCCTATACTTT CTAGAG | 70.08 | This work |
| Iend-Km For2 | TTAATTAAAGATGTGTATAAGAGACAG | 55.9 | This work |
| Tnext(Tel) | CGACCTGTTGGTGATGGAG | 62.32 | This work |
| Iend F | AGATCTGATCAAGAGACAGT | 56.3 | This work |

**Table 8-3: List of primers used for the Southern blot experiments**

| Primers | Sequence (5'-3') | Melting Temperature (°C) | Reference |
|---|---|---|---|
| Km-pBAM F | TATTCAGCGTGAAACGAGCTGTAG | 61.0 | Kindly provided by S. Arias Rivas |
| Km-pBAM R | CGGATTATCAATGCCATATTTCTG | 57.6 | Kindly provided by S. Arias Rivas |
| Tel fw | GAAGCAGGCGAGAAACTGAC | 59.4 | This work |
| Tel rv | TTGATGAGCGTCGTCTGAAC | 57.3 | This work |

**Table 8-4: List of primers used for the verification of the *P. putida* mutants**

| Primers | Sequence (5'-3') | Melting Temperature (°C) | Reference |
|---|---|---|---|
| PP_0072 F | GAGATGGCAGAGGGTGTGAT | 59.4 | This work |
| PP_0072 R | GTGAGCTCTGCCGAGAAGG | 61.0 | This work |
| PP_0245 F | GTTGGTCACCACGCCTTCG | 61.0 | This work |
| PP_0245 R | GCGGTAGTCGAGGACTGTGT | 61.4 | This work |
| PP_0597 F | CAGGAACTGGACCGCTTG | 58.2 | This work |
| PP_0597 R | AGTTGAATGGGGTGATACCG | 57.3 | This work |
| PP_0659 F | CTTCATCTACAGCCGCATGA | 57.3 | This work |
| PP_0659 R | GCGCTCTTCCAAAGTGTTCT | 57.3 | This work |
| PP_1626 F | CTTCTCGAAGCGGTAGCTGT | 59.4 | This work |
| PP_1626 R | GGAAGCTGAAAAACCAGCAC | 57.3 | This work |
| PP_1993 F | CCTGTCGTTGGAAGACGATT | 57.3 | This work |
| PP_1993 R | CGGGGATTATAACGACAACG | 57.3 | This work |
| PP_3490 F | GCTCGGCATACAGGGTGTAG | 61.4 | This work |
| PP_3490 R2 | TTCTTCTGGCACACACCTTG | 57.3 | This work |
| PP_3495 F2 | AAAACAAGGAGCACCCCTTC | 57.3 | This work |
| PP_3529 F | ACCTACGAACACCCCTACCC | 61.4 | This work |
| PP_3529 R | ACTACATTGGCCAGGTCGTC | 59.4 | This work |

| | | | |
|---|---|---|---|
| PP_3511 F | GGCACTTGCAAGACAATGAG | 57.3 | This work |
| PP_3511 R | CGTTTCAGGCGACCTTGACG | 61.4 | This work |
| PP_3512 F3 | AGCCAATAATGCCTGGGAGAT | 57.9 | This work |
| PP_3512 R3 | CCTCGAGAATTGCTTCGACAT | 57.9 | This work |
| PP_3534 F | GGATGCCACTGAGGCTAGAG | 61.4 | This work |
| PP_3733 F | ATAGTACGAAGCCTGGCCCT | 57.3 | This work |
| PP_3733 R2 | AAGCTTAGTGCGTGGCTGTT | 59.4 | This work |
| Oend-Km For2 | TTTCACAAAACGGTTTACAAGCATA | 56.4 | Kindly provided by S. Arias Rivas |

## 8.2 *P. putida* SMT mutants library

**Table 8-5: Genes and intergenic regions hit by a mini-Tn*5 KpF* in *P. putida* SMT single mutants**

| Col. S | Locus ID | *gene* | Sign | Product name | K/EC numbers | **Cellular role category** |
|---|---|---|---|---|---|---|
| 109 112 | PP_0051 | - | (+) | sigma-54 dependent transcriptional regulator | - | **Regulatory functions:** Protein interactions |
| 10 | PP_0072 | *qor-1* | (-) | quinone oxidoreductase | K00344 NADPH2:quinone reductase [EC:1.6.5.5] | **Energy metabolism:** Electron transport |
| 9 | PP_0077 | *betC* | (+) | choline sulfatase | choline-sulfatase [EC:3.1.6.6] | **Central intermediary metabolism:** Sulfur metabolism |
| 16 | PP_0168 | *lapA* | (+) | large adhesion, surface associated. Biofilm formation | surface adhesion protein | **Cellular processes:** Cell adhesion |
| 39 | PP_0266 | - | (-) | agmatine deiminase | K10536 agmatine deiminase [EC:3.5.3.12] | **Central intermediary metabolism:** Polyamine biosynthesis |
| 44 | PP_0339 | *aceE* | (-) | pyruvate dehydrogenase subunit E1 | K00163 pyruvate dehydrogenase E1 component [EC:1.2.4.1] | **Energy metabolism:** Pyruvate dehydrogenase |
| 11 | PP_0350 | - | (+) | TonB-dependent siderophore receptor | K02014 iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| 45 | PP_0405 | - | (-) | aminoglycoside phosphotransferase | K07102 | **Cellular processes:** Toxin production and resistance |
| 18 | IR_0780 -81 | - | (-) | hypothetical protein / hypothetical protein | - | - |
| 35 | PP_0789 | *ampD* | (-) | N-acetyl-anhydromuranmyl-L-alanine amidase | AmpD protein | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides / **Cellular processes:** Toxin production and resistance |
| 99 | PP_0806 | *lapF* | (-) | seed colonization adhesion protein LapF | - | **Cellular processes:** Cell adhesion |
| 89 | PP_0813 | *cyoB* | (+) | cytochrome o ubiquinol oxidase, subunit I | cytochrome o ubiquinol oxidase subunit I [EC:1.10.3.-] | **Energy metabolism:** electron transport |
| 22 | PP_0886 | - | (+) | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |

| | | | | | | |
|---|---|---|---|---|---|---|
| 32 | PP_0896 | - | (-) | nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase | - | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides / **Fatty acid and phospholipid metabolism:** Biosynthesis |
| 5 | PP_1490 | - | (-) | methyltransferase, CheR-like | chemotaxis protein methyltransferase CheR [EC:2.1.1.80] | **Cellular processes:** chemotaxis and motility |
| 14 | PP_1690 | - | (-) | conserved hypothetical protein | - | **Hypothetical proteins:** conserved |
| 2 | PP_1695 | - | (-) | sodium-solute symporter/sensory box histidine kinase/response regulator, putative | - | **Transport and binding proteins:** Unknown substrate / **Regulatory functions:** Small molecule interactions / **Signal transduction:** Two-component systems |
| 6 | PP_1869 | - | (-) | conserved hypothetical protein | - | **Hypothetical proteins:** conserved |
| 46 88 | PP_1963 | - | (-) | hypothetical protein | - | - |
| 12 | IR_2118 -19 | - | (-) | hypothetical protein/ lipid ABC transporter ATPase/inner membrane protein | /K06147 ATP-binding cassette, subfamily B, bacterial | **Hypothetical proteins:** Conserved/ **Transport and binding proteins:** Unknown substrate |
| 38 | PP_2122 | *moaB-1* | (-) | molybdenum cofactor biosynthesis protein B | molybdenum cofactor biosynthesis protein B | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Molybdopterin |
| 25 | PP_2402 | - | (-) | integral membrane sensor signal transduction histidine kinase | - | **Regulatory functions:** Protein interactions / **Signal transduction:** Two-component systems |
| 98 | PP_2820 | *nfxB* | (+) | transcriptional regulator NfxB | - | **Cellular processes:** Detoxification / **Regulatory functions:** DNA interactions |
| 15 | PP_2838 | - | (+) | VRR-NUC domain protein | - | **Unknown function:** general |
| 23 24 | PP_3084 | - | (+) | TonB-dependent siderophore receptor | iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| 20 | PP_3132 | - | (-) | polysaccharide biosynthesis protein | - | **Transport and binding proteins:** Other |
| 37 | PP_3210 | - | (+) | ABC transporter, permease protein | K02050 sulfonate/nitrate/taurine transport system permease protein | **Transport and binding proteins:** Unknown substrate |
| 36 | PP_3461 | - | (+) | radical SAM domain protein | K06871 | **Unknown function:** Enzymes of unknown specificity |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3<br>4 | PP_3490 | - | (+) | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| 7<br>17 | PP_3534 | - | (+) | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| 27 | PP_3545 | - | (+) | PAS/PAC sensor hybrid histidine kinase | - | **Regulatory functions:** Small molecule interactions **/ Signal transduction:** Two-component systems |
| 33 | PP_3820 | - | (+) | group II intron-encoding maturase | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| 19 | PP_3913 | - | (-) | hypothetical protein | - | - |
| 28 | PP_3914 | - | (-) | hypothetical protein | - | - |
| 31 | PP_4050 | *glgA* | (+) | glycogen synthase (EC:2.4.1.21) | starch synthase [EC:2.4.1.21] | **Energy metabolism:** Biosynthesis and degradation of polysaccharides |
| 26 | PP_4056 | - | (+) | endonuclease/exonuclease/phosphatase | K06896 | **DNA metabolism:** DNA replication, recombination, and repair **/ Unknown function:** Enzymes of unknown specificity |
| 102<br>104<br>105 | PP_4121 | *nuoCD* | (-) | bifunctional NADH:ubiquinone oxidoreductase subunit C/D (EC:1.6.99.5) | K00333 NADH dehydrogenase I subunit C/D [EC:1.6.5.3] | **Energy metabolism:** Electron transport |
| 40 | PP_4124 | *nuoG* | (-) | NADH dehydrogenase subunit G | NADH dehydrogenase I subunit G [EC:1.6.5.3] | **Energy mtabolism:** electron transport |
| 48 | PP_4185 | *sucD* | (+) | succinyl-CoA synthetase subunit alpha | succinyl-CoA synthetase alpha subunit [EC:6.2.1.5] | **Energy metabolism:** TCA cycle |
| 21 | PP_4231 | - | (+) | xanthine dehydrogenase accessory factor, putative | xanthine dehydrogenase accessory factor | **Purines, pyrimidines, nucleosides, and nucleotides:** Other |
| 29 | PP_4525 | - | (+) | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| 96 | PP_4620 | - | (-) | fumarylacetoacetase | fumarylacetoacetase [EC:3.7.1.2] | **Energy metabolism:** Amino acids and amines |
| 30 | PP_4686 | - | (-) | lipoprotein, putative | - | **Cell envelope:** Other |
| 101 | PP_4754 | - | (-) | TetR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |

| | | | | | | |
|---|---|---|---|---|---|---|
| 100 | PP_4761 | - | (-) | HAD superfamily hydrolase | - | **Unknown function:** Enzymes of unknown specificity |
| 8 | PP_4828 | *cobH* | (-) | precorrin-8X methylmutase [EC:5.4.1.2] | precorrin-8X methylmutase [EC:5.4.1.2] | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Heme, porphyrin, and cobalamin |
| 92 | PP_4971 | - | (-) | MltA domain protein | membrane-bound lytic murein transglycosylase A [EC:3.2.1.-] | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides |
| 1 | PP_5025 | *mdoH* | (+) | glucosyltransferase MdoH | membrane glycosyltransferase [EC:2.4.1.-] | **Cell envelope: B**iosynthesis and degradation of surface polysaccharides and lipopolysaccharides |
| 42 | PP_5061 | - | (+) | choline/carnitine/betaine transporter family protein | high-affinity choline transport protein | **Transport and binding proteins:** Other |
| 41 | PP_5063 | *betB* | (+) | betaine aldehyde dehydrogenase | betaine-aldehyde dehydrogenase [EC:1.2.1.8] | **Energy metabolism:** fermentation |
| 43 | PP_5064 | *betA* | (+) | choline dehydrogenase | choline dehydrogenase [EC:1.1.99.1] | **Energy metabolism:** fermentation |
| 110 | PP_5218 | - | (-) | DedA family protein | - | **Unknown function:** General |
| 93 | PP_5246 | *kefB* | (-) | potassium efflux system protein | glutathione-regulated potassium-efflux system protein KefB | **Transport and binding proteins:** Cations and iron carrying compounds |
| 13 | IR_5338 -39 | *aspA/* | (+) | aspartate ammonia-lyase/ AraC family transcriptional regulator | aspartate ammonia-lyase [EC:4.3.1.1]/ | **Energy metabolism:** Amino acids and amines/ **Regulatory functions:** DNA interactions |

After triparental mating, the pBAM1/*KpF* vector was transferred to the *P. putida* TEC1 wild-type strain. The 54 independent genes or intergenic regions (IR), hit by a mini-Tn*5 KpF*, were presented in the Table 8-5. The first column indicated for which *P. putida* SMT colony the gene or intergenic region was described. The abbreviation used for naming the colonies of *P. putida* SMT strains was col. S, however only the numbers were reported in each row. Information about the Locus ID (PP), the product name, the orientation sign, as well as the K and/or EC number and the cellular role category were provided. The orientation sign either positive (‘+’) or negative (‘-’) indicated the orientation of the *FRT* site relatively to the ‘+’ and ‘-’ strands of the genome; therefore the orientation of the mini-Tn*5 KpF* in the chromosome. For a better reading of the table a single color was associated to a specific cellular role category. The different colors used in the table were summarized hereafter.

| | |
|---|---|
| 2 | Biosynthesis of cofactors, prosthetic groups and carriers |
| 5 | Cell envelope |
| 6 | Cellular processes |
| 2 | Central intermediary metabolism |
| 1 | DNA metabolism |
| 8 | Energy metabolism |
| 1 | Fatty acid and phospholipid metabolism |
| 1 | Mobile and extrachromosomal element functions |
| 1 | Purines, pyrimidines, nucleosides, and nucleotides |
| 8 | Regulatory functions |
| 3 | Signal transduction |
| 7 | Transport and binding proteins |

The numbers indicated in the colored rectangles corresponded to the amount of apparition for each role category. When a mini-Tn*5 KpF* hit an intergenic region, the description for both the flanking genes was given in the same row, but the cellular role category was not taken in account for establishing the previous numbers.

The genes encoding conserved hypothetical proteins were not associated with any color, therefore left in white. The proteins with unknown functions and proteins for which no role was described in any database were as well left in white.

## 8.3 *P. putida* TMT mutants

**Table 8-6: Configuration of the mini-Tn*5 KpF* and mini-Tn*5 TF* for the *P. putida* TMT mutant strains**

| Col. T | **Position mini-Tn*5 KpF*** | Orientation | Detection Oend | **Position mini-Tn*5 TF*** | Orientation | Detection Oend | Putative deleted fragment |
|---|---|---|---|---|---|---|---|
| 1 | **PP_0072** | (-) | yes | **PP_1513** | (-) | yes | ~1735k b |
| 2 | **PP_0072** | (-) | yes | **PP_4031** | (-) | yes | ~1722 kb |
| 3 | **PP_0072** | (-) | yes | **PP_3589** | (-) | yes | ~2185 kb |
| 5 | **PP_1869** | (-) | yes | **PP_3869** | (+) | yes | - |
| 6 | **PP_5025** | (+) | no | **PP_5093** | (-) | yes | - |
| 7 | **PP_5025** | (+) | yes | **PP_4224** | (-) | yes | - |
| 9 | **PP_1490** | (-) | yes | **PP_0873** | (+) | yes | - |
| 12 | **PP_0072** | (-) | yes | **PP_3364** | (-) | yes | ~2457 kb |
| 13 | **PP_0072** | (-) | yes | **PP_2534** | (+) | yes | - |
| 14 | **PP_0072** | (-) | yes | **PP_1075** | (+) | no | - |
| 21 | **PP_0077** | (+) | yes | **PP_0168** | (+) | yes | ~100 kb |
| 22 | **PP_3490** | (+) | yes | **PP_2404** | (-) | yes | - |
| 24 | **PP_3534** | (+) | no | **PP_2695** | (-) | yes | - |
| 27 | **PP_0072** | (-) | yes | **PP_4111** | (+) | yes | - |
| 34 | **PP_1490** | (-) | yes | **PP_3260** | (-) | no | ~2000 kb |
| 35 | **PP_0077** | (+) | yes | **PP_4029** | (-) | no | - |
| 37 | **PP_3490** | (+) | yes | **PP_5188** | (+) | yes | ~1959 kb |
| 38 | **PP_1490** | (-) | yes | **PP_5366** | (+) | yes | - |
| 40 | **PP_5025** | (+) | yes | **PP_4316** | (-) | yes | - |
| 51 | **PP_1869** | (-) | yes | **PP_5241** | (+) | yes | - |
| 53 | **PP_1869** | (-) | no | **PP_1896** | (-) | yes | ~46 kb |
| 54 | **PP_0072** | (-) | yes | **PP_2356** | (+) | yes | - |
| 54 | **PP_0072** | (-) | yes | **PP_2356** | (+) | yes | - |
| 57 | **PP_0077** | (+) | yes | **PP_1776** | (-) | yes | - |
| 58 | **PP_1490** | (-) | yes | **PP_3827** | (-) | yes | ~2658 kb |
| 59 | **PP_5025** | (+) | yes | **PP_0902** | (+) | yes | ~1500 kb |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 60 | **PP_5025** | (+) | yes | **PP_2870** | (+) | yes | ~2452 kb |
| 61 | **PP_0072** | (-) | no | **PP_2588** | (-) | yes | ~2874 kb |
| 62 | **PP_1869** | (-) | yes | **PP_3490** | (-) | yes | ~1865 kb |
| 63 | **PP_3534** | (+) | yes | **PP_2561** | (-) | yes | - |
| 64 | **PP_0072** | (-) | yes | **PP_2281** | (-) | yes | ~2520 kb |
| 67 | **PP_1869** | (-) | yes | **PP_1952** | (-) | yes | ~117 kb |
| 69 | **PP_0072** | (+) | yes | **PP_0179** | (+) | yes | ~149 kb |
| 70 | **PP_1490** | (-) | yes | **PP_1740** | (+) | yes | - |
| 71 | **PP_0077** | (+) | yes | **PP_2991** | (-) | yes | - |
| 72 | **PP_3490** | (+) | yes | **PP_0370** | (+) | yes | ~2671 kb |
| 73 | **PP_0072** | (-) | yes | **PP_2841** | (+) | yes | - |
| 74 | **PP_0072** | (-) | yes | **PP_1645** | (-) | yes | ~1757 kb |
| 75 | **PP_0072** | (-) | yes | **PP_5261** | (-) | yes | ~256 kb |
| 76 | **PP_1490** | (-) | yes | **PP_1141** | (+) | yes | - |
| 79 | **PP_0072** | (+) | yes | **PP_0806** | (+) | no | ~850 kb |
| 82 | **PP_3490** | (+) | yes | **PP_3262** | (+) | yes | ~263 kb |
| 83 | **PP_1695** | (-) | yes | **PP_3498** | (-) | yes | ~2081 kb |
| 85 | **PP_1869** | (-) | yes | **PP_5067** | (-) | yes | ~2492 kb |
| 86 | **PP_1490** | (-) | yes | **PP_2900** | (-) | yes | ~1603 kb |
| 88 | **PP_0072** | (-) | yes | **PP_4688** | (-) | yes | ~935 kb |
| 90 | **PP_1490** | (-) | yes | **IR_4446-47** | (+) | yes | - |
| 91 | **PP_3490** | (+) | yes | **PP_3529** | (+) | yes | ~41.5 kb |
| 94 | **PP_1869** | (-) | yes | **PP_3915** | (+) | yes | - |
| 95 | **PP_5025** | (+) | yes | **PP_2032** | (-) | yes | - |
| 97 | **PP_0072** | (-) | yes | **PP_1449** | (-) | yes | ~1570 kb |
| 99 | **PP_3534** | (+) | yes | **PP_2696** | (-) | yes | - |
| 101 | **PP_1869** | (-) | yes | **PP_5405** | (+) | yes | - |
| 102 | **PP_5025** | (+) | yes | **PP_0167** | (-) | yes | - |
| 103 | **PP_1695** | (-) | yes | **PP_3329** | (-) | yes | ~1880 kb |
| 104 | **PP_1869** | (-) | yes | **PP_0971** | (+) | yes | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 105 | **PP_0072** | (-) | yes | **PP_5261** | (-) | yes | - |
| 106 | **PP_1490** | (-) | yes | **PP_1281** | (+) | yes | - |
| 107 | **PP_3534** | (+) | no | **PP_0323** | (-) | yes | - |
| 113 | **PP_0077** | (+) | yes | **PP_0971** | (+) | yes | ~1023 kb |
| 115 | **PP_0072** | (-) | yes | **PP_4946** | (-) | yes | - |
| 116 | **PP_1869** | (-) | yes | **IR_2994-95** | (-) | no | ~1299 kb |
| 119 | **PP_1869** | (-) | yes | **PP_0976** | (-) | yes | ~978 kb |
| 120 | **PP_0072** | (-) | yes | **PP_3900** | (-) | no | ~1849 kb |
| 123 | **PP_1869** | (-) | yes | **PP_0224** | (+) | yes | - |
| 126 | **PP_5025** | (+) | not sure | **PP_2762** | (+) | yes | ~2578 kb |
| 132 | **PP_3490** | (+) | yes | **IR_3260-61** | (-) | yes | - |
| 136 | **PP_1869** | (-) | yes | **PP_2703** | (+) | yes | - |
| 141 | **PP_4828** | (-) | yes | **PP_0489** | (-) | no | ~1263 kb |
| 144 | **PP_3490** | (+) | yes | **PP_4455** | (-) | yes | - |
| 145 | **PP_3534** | (+) | yes | **PP_1277** | (-) | yes | - |
| 146 | **PP_1490** | (-) | yes | **PP_2395** | (+) | yes | - |
| 147 | **PP_1869** | (-) | yes | **PP_3589** | (-) | yes | ~1985 kb |
| 148 | **PP_0072** | (-) | yes | **PP_3279** | (+) | yes | - |
| 149 | **PP_0072** | (-) | yes | **PP_3988** | (+) | no | - |
| 150 | **PP_3490** | (+) | yes | **PP_0234** | (-) | yes | - |
| 155 | **PP_1869** | (-) | yes | **PP_2209** | (+) | yes | - |
| 157 | **PP_0072** | (-) | yes | **PP_2558** | (+) | yes | - |
| 166 | **PP_1490** | (-) | yes | **PP_3796** | (+) | no | - |
| 171 | **PP_3490** | (+) | yes | **PP_5154** | (-) | no | - |
| 197 | **PP_0077** | (+) | yes | **PP_1866** | (+) | yes | ~2000 kb |
| 198 | **PP_1869** | (-) | yes | **IR_3795-96** | (+) | yes | - |
| 199 | **PP_0072** | (-) | yes | **PP_2286** | (-) | yes | - |
| 200 | **PP_0072** | (-) | yes | **PP_4830** | (+) | yes | - |
| 201 | **PP_0072** | (-) | yes | **PP_2600** | (-) | yes | ~2890 kb |
| 203 | **PP_3534** | (+) | yes | **PP_0672** | (-) | yes | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 204 | **PP_1490** | (-) | yes | **PP_2369** | (-) | yes | ~1010 kb |
| 205 | **PP_0077** | (+) | yes | **IR_3153-52** | (-) | yes | - |
| 207 | **PP_0077** | (+) | yes | **PP_3539** | (-) | yes | - |
| 208 | **PP_0072** | (-) | yes | **PP_2179** | (-) | yes | ~2403 kb |
| 209 | **PP_0072** | (-) | yes | **PP_2642** | (-) | no | ~2945 kb |
| 211 | **PP_1695** | (-) | yes | **IR_5339-40** | (-) | yes | ~1982 kb |
| 214 | **PP_1869** | (-) | yes | **PP_4330** | (-) | no | ~2829 kb |
| 215 | **PP_0072** | (-) | yes | **PP_2542** | (-) | no | ~2805 kb |
| 216 | **PP_0072** | (-) | yes | **PP_2787** | (+) | yes | - |
| 219 | **PP_3534** | (+) | yes | **PP_3124** | (-) | no | - |
| 220 | **PP_0072** | (-) | yes | **PP_3124** | (-) | yes | ~2727 kb |
| 227 | **PP_3490** | (+) | yes | **PP_1072** | (-) | no | - |
| 228 | **PP_3490** | (+) | yes | **PP_0213** | (+) | yes | ~2485 kb |
| 232 | **PP_3490** | (+) | yes | **PP_3315** | (+) | yes | ~209 kb |
| 233 | **PP_0072** | (-) | yes | **PP_2749** | (-) | yes | - |
| 241 | **PP_0072** | (-) | yes | **IR_1871-72** | (+) | yes | - |
| 248 | **PP_1695** | (-) | yes | **IR_2541-42** | (+) | yes | - |
| 251 | **PP_3490** | (+) | yes | **PP_5334** | (+) | no | ~2122 kb |
| 252 | **PP_3490** | (+) | yes | **PP_4683** | (-) | yes | - |
| 253 | **PP_3490** | (+) | yes | **PP_4011** | (+) | no | ~561 kb |
| 257 | **PP_0072** | (-) | yes | **PP_0918** | (+) | no | - |
| 259 | **PP_1490** | (-) | yes | **PP_2650** | (+) | no | - |
| 260 | **PP_3534** | (+) | yes | **PP_0486** | (-) | yes | - |
| 265 | **PP_0072** | (-) | yes | **PP_5318** | (+) | no | - |
| 295 | **PP_5025** | (+) | yes | **PP_1831** | (+) | no | ~2500 kb |
| 297 | **PP_0072** | (-) | yes | **PP_1127** | (-) | yes | ~1206 kb |
| 301 | **PP_0072** | (-) | yes | **PP_1308** | (-) | yes | ~1413 kb |
| 302 | **PP_3490** | (+) | yes | **PP_2859** | (-) | no | - |
| 308 | **PP_4828** | (-) | yes | **IR_2118-19** | (-) | no | ~3104 kb |
| 366 | **PP_1869** | (-) | yes | **PP_0913** | (+) | yes | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 368 | **PP_5025** | (+) | yes | **PP_0065** | (-) | yes | - |
| 374 | **PP_0072** | (-) | yes | **PP_0913** | (+) | yes | - |
| 376 | **PP_5025** | (+) | yes | **PP_4954** | (+) | yes | ~77 kb |
| 378 | **PP_3490** | (+) | yes | **PP_2270** | (-) | no | - |
| 379 | **PP_0072** | (-) | yes | **PP_1105** | (+) | yes | - |
| 380 | **PP_1695** | (-) | yes | **PP_0043** | (-) | yes | ~1840 kb |
| 381 | **PP_0072** | (-) | yes | **PP_1018** | (-) | no | ~1079 kb |
| 382 | **PP_0072** | (-) | yes | **PP_5026** | (+) | yes | - |
| 385 | **PP_0072** | (-) | yes | **PP_3396** | (+) | yes | - |
| 386 | **PP_3490** | (+) | yes | **PP_4223** | (-) | yes | - |
| 389 | **PP_0072** | (-) | yes | **PP_1816** | (+) | yes | - |
| 390 | **PP_0072** | (-) | yes | **PP_2576** | (-) | no | ~2860 kb |
| 394 | **PP_1869** | (-) | yes | **IR_2702-03** | (+) | yes | - |
| 395 | **PP_3490** | (+) | yes | **PP_5252** | (+) | yes | ~2037 kb |
| 397 | **PP_1695** | (-) | yes | **PP_2590** | (-) | no | ~1073 kb |
| 398 | **PP_5025** | (+) | yes | **PP_0738** | (-) | yes | - |
| 399 | **PP_3490** | (+) | yes | **PP_1298** | (-) | yes | - |
| 400 | **PP_3490** | (+) | yes | **PP_0612** | (+) | no | ~2943 kb |
| 404 | **PP_3534** | (+) | yes | **PP_2645** | (+) | yes | ~974 kb |
| 406 | **PP_0072** | (-) | yes | **PP_3733** | (+) | yes | - |
| 407 | **PP_3534** | (+) | yes | **PP_3733** | (+) | yes | ~254 kb |
| 409 | **PP_3490** | (+) | yes | **PP_0825** | (-) | no | - |
| 410 | **PP_1490** | (-) | yes | **PP_5372** | (+) | yes | - |
| 413 | **PP_1695** | (-) | yes | **PP_4358** | (-) | no | ~3062 kb |
| 414 | **PP_0072** | (-) | yes | **PP_0958** | (-) | yes | ~1017 kb |
| 416 | **PP_3534** | (+) | yes | **PP_0411** | (-) | yes | - |
| 418 | **PP_0077** | (+) | yes | **PP_2073** | (+) | yes | ~2270 kb |
| 419 | **PP_1490** | (-) | yes | **PP_4736** | (-) | yes | ~2490 kb |
| 420 | **PP_3490** | (+) | yes | **PP_3327** | (+) | yes | ~196 kb |
| 423 | **PP_4828** | (-) | yes | **PP_2822** | (+) | yes | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 425 | **PP_5025** | (+) | yes | **PP_0716** | (+) | yes | ~1290 kb |
| 427 | **PP_0072** | (-) | yes | **PP_4468** | (+) | no | - |
| 428 | **PP_0072** | (-) | yes | **PP_3660** | (-) | yes | ~2106 kb |
| 430 | **PP_3534** | (+) | yes | **PP_3733** | (+) | yes | ~254 kb |
| 431 | **PP_5025** | (+) | yes | **PP_3483** | (+) | yes | ~1777 kb |
| 432 | **PP_1695** | (-) | yes | **PP_0928** | (+) | yes | - |
| 433 | **PP_0072** | (-) | yes | **PP_2409** | (+) | yes | - |
| 434 | **PP_3534** | (+) | yes | **PP_0371** | (+) | yes | ~2628 kb |
| 436 | **PP_1869** | (-) | yes | **PP_2080** | (+) | yes | - |
| 438 | **PP_3490** | (+) | yes | **PP_4916** | (-) | yes | - |
| 439 | **PP_0072** | (-) | yes | **PP_4019** | (+) | yes | - |
| 441 | **PP_0072** | (-) | yes | **PP_3660** | (-) | no | ~2106 kb |
| 442 | **PP_3534** | (+) | yes | **PP_0325** | (-) | no | - |
| 443 | **PP_3534** | (+) | yes | **PP_3597** | (-) | yes | - |
| 444 | **PP_3490** | (+) | yes | **IR_0655-56** | (+) | yes | ~2985 kb |
| 445 | **PP_1869** | (-) | no | **PP_1072** | (-) | yes | ~863 kb |
| 446 | **PP_4828** | (-) | yes | **PP_0887** | (+) | no | - |
| 447 | **PP_5025** | (+) | yes | **PP_2914** | (-) | yes | - |
| 448 | **PP_0072** | (-) | yes | **PP_2554** | (+) | yes | - |
| 449 | **PP_5025** | (+) | yes | **PP_1154** | (-) | yes | - |
| 451 | **PP_3490** | (+) | yes | **PP_0302** | (-) | yes | - |
| 452 | **PP_5025** | (+) | yes | **PP_1838** | (-) | yes | - |
| 454 | **PP_1695** | (-) | yes | **PP_3224** | (+) | yes | - |
| 456 | **PP_0072** | (-) | yes | **IR_2348-49** | (+) | yes | - |
| 457 | **PP_1695** | (-) | yes | **PP_1626** | (+) | yes | - |
| 458 | **PP_3534** | (+) | yes | **PP_3733** | (+) | yes | ~254 kb |
| 460 | **PP_0072** | (-) | yes | **PP_4626** | (+) | no | - |
| 462 | **PP_0072** | (-) | yes | **PP_0573** | (+) | no | - |
| 464 | **PP_0072** | (-) | yes | **PP_1055** | (+) | no | - |
| 466 | **PP_1695** | (-) | yes | **PP_0768** | (+) | yes | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 467 | **PP_3534** | (+) | yes | **PP_0483** | (+) | yes | ~2182 kb |
| 468 | **PP_1869** | (-) | yes | **PP_0041** | (+) | yes | - |
| 471 | **PP_0072** | (-) | yes | **PP_1758** | (+) | no | - |
| 473 | **PP_0072** | (-) | yes | **IR_3603-04** | (-) | yes | ~2168 kb |
| 475 | **PP_0072** | (-) | yes | **IR_5275-76** | (-) | no | ~239 kb |
| 477 | **PP_3534** | (+) | yes | **PP_3074** | (+) | yes | ~547 kb |
| 478 | **PP_5025** | (+) | yes | **PP_0144** | (+) | yes | ~610 kb |
| 479 | **PP_1695** | (-) | yes | **PP_5009** | (+) | yes | - |
| 481 | **PP_3490** | (+) | yes | **PP_1283** | (+) | no | ~2492 kb |
| 482 | **PP_3490** | (+) | yes | **PP_3578** | (-) | yes | - |
| 486 | **PP_3490** | (+) | yes | **IR_4979-80** | (+) | no | ~1713 kb |
| 487 | **PP_1869** | (-) | yes | **PP_2180** | (-) | no | ~392 kb |
| 489 | **PP_5025** | (+) | yes | **PP_0348** | (+) | no | ~879 kb |
| 490 | **PP_3534** | (+) | yes | **PP_2717** | (-) | yes | - |
| 491 | **PP_0072** | (-) | yes | **PP_5192** | (-) | yes | ~340 kb |
| 492 | **PP_4828** | (-) | yes | **PP_4516** | (+) | yes | - |
| 493 | **PP_3490** | (+) | yes | **PP_4075** | (+) | yes | ~644 kb |
| 494 | **PP_0072** | (-) | yes | **PP_4379** | (+) | no | - |
| 497 | **PP_1869** | (-) | yes | **PP_4918** | (+) | yes | - |
| 498 | **PP_1869** | (-) | yes | **PP_0325** | (-) | no | ~1701 kb |
| 499 | **PP_3490** | (+) | yes | **PP_3970** | (-) | yes | - |
| 501 | **PP_3490** | (+) | yes | **PP_0350** | (-) | yes | - |
| 506 | **PP_1490** | (-) | yes | **PP_1740** | (+) | yes | - |
| 507 | **PP_0072** | (-) | yes | **PP_5045** | (+) | yes | - |
| 508 | **PP_3490** | (+) | yes | **PP_4470** | (-) | yes | - |
| 509 | **PP_1695** | (-) | yes | **PP_3583** | (-) | yes | ~2184 kb |
| 510 | **PP_3490** | (+) | yes | **PP_0348** | (+) | yes | ~2645 kb |
| 513 | **PP_3490** | (+) | yes | **PP_2613** | (-) | yes | - |
| 514 | **PP_0072** | (-) | yes | **PP_3786** | (-) | yes | ~1949 kb |
| 515 | **PP_0072** | (-) | yes | **PP_0287** | (+) | no | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 519 | **PP_3490** | (+) | yes | **PP_3450** | (-) | no | - |
| 520 | **PP_0072** | (-) | yes | **PP_3498** | (-) | yes | ~2295 kb |
| 521 | **PP_3490** | (+) | yes | **IR_3273-74** | (+) | no | ~251 kb |
| 523 | **PP_1869** | (-) | yes | **PP_2109** | (+) | yes | - |
| 524 | **PP_1869** | (-) | yes | **PP_4852** | (-) | yes | ~2755 kb |
| 525 | **PP_1869** | (-) | yes | **PP_0583** | (+) | no | - |
| 526 | **PP_5025** | (+) | yes | **IR_1036-37** | (-) | yes | - |
| 527 | **PP_4828** | (-) | yes | **PP_4278** | (+) | yes | - |
| 528 | **PP_0072** | (-) | yes | **PP_4217** | (-) | yes | ~1498 kb |
| 530 | **PP_1490** | (-) | yes | **PP_1751** | (-) | yes | ~ 258 kb |
| 533 | **PP_1869** | (-) | yes | **PP_3846** | (-) | yes | ~2276 kb |
| 534 | **PP_0072** | (-) | yes | **PP_0245** | (-) | yes | ~ 217 kb |
| 535 | **PP_1869** | (-) | yes | **IR_4467-68** | (-) | yes | ~2982 kb |
| 536 | **PP_0072** | (-) | yes | **PP_4610** | (-) | yes | ~1031 kb |
| 537 | **PP_3534** | (+) | yes | **PP_2144** | (+) | yes | ~1560 kb |
| 539 | **PP_0072** | (-) | yes | **IR_5172-73** | (-) | yes | ~366 kb |
| 540 | **PP_0072** | (-) | yes | **PP_4146** | (+) | yes | - |
| 541 | **PP_0072** | (-) | yes | **PP_4146** | (+) | yes | - |
| 542 | **PP_0072** | (-) | yes | **IR_5347-48** | (+) | yes | - |
| 543 | **PP_3490** | (+) | yes | **PP_1648** | (-) | yes | - |
| 545 | **PP_0072** | (-) | yes | **PP_2444** | (+) | yes | - |
| 549 | **PP_0072** | (-) | yes | **PP_3038** | (-) | yes | ~2830 kb |
| 550 | **PP_0072** | (-) | yes | **PP_0597** | (-) | yes | ~613 kb |
| 552 | **PP_1695** | (-) | yes | **IR_4771-72** | (+) | yes | - |
| 553 | **PP_3490** | (+) | yes | **IR_3232-33** | (+) | yes | ~290 kb |
| 556 | **PP_0072** | (-) | yes | **PP_3625** | (+) | yes | - |
| 558 | **PP_1869** | (-) | yes | **PP_4441** | (-) | yes | ~2943 kb |
| 560 | **PP_0072** | (-) | no | **PP_0499** | (-) | no | ~506 kb |
| 566 | **PP_1490** | (-) | yes | **PP_5068** | (+) | yes | - |
| 569 | **PP_1695** | (-) | yes | **PP_0593** | (-) | no | ~1195 kb |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 573 | **PP_3490** | (+) | yes | **PP_3092** | (-) | no | - |
| 576 | **PP_3490** | (+) | yes | **PP_4487** | (-) | yes | - |
| 577 | **PP_1869** | (-) | yes | **PP_4644** | (+) | no | - |
| 580 | **PP_0072** | (-) | yes | **PP_4475** | (-) | no | ~1179 kb |
| 583 | **PP_1490** | (-) | yes | **PP_2453** | (+) | yes | - |
| 584 | **PP_0072** | (-) | yes | **PP_1983** | (-) | no | ~2165 kb |
| 585 | **PP_1869** | (-) | yes | **PP_1225** | (-) | yes | ~691 kb |
| 588 | **PP_0072** | (-) | yes | **PP_2334** | (+) | yes | - |
| 599 | **PP_3490** | (+) | yes | **PP_0207** | (-) | no | - |
| 602 | **PP_1869** | (-) | yes | **PP_1385** | (-) | yes | ~513 kb |
| 603 | **PP_1869** | (-) | yes | **PP_2677** | (-) | yes | ~975 kb |
| 604 | **PP_0072** | (-) | yes | **PP_0235** | (+) | yes | - |
| 605 | **PP_5025** | (+) | yes | **PP_3514** | (+) | yes | ~1737 kb |
| 606 | **PP_3490** | (+) | yes | **PP_3251** | (-) | no | - |
| 607 | **PP_3490** | (+) | yes | **PP_1090** | (+) | no | ~2710 kb |
| 608 | **PP_0072** | (-) | yes | **PP_4220** | (+) | no | - |
| 611 | **PP_3490** | (+) | yes | **PP_4683** | (+) | yes | ~1362 kb |
| 612 | **PP_0077** | (+) | yes | **PP_2853** | (+) | yes | ~3013 kb |

The *P. putida* TMT mutant strains, for which both mini-Tn*5* derivatives were mapped, were reported in Table 8-6. In total 250 double knock-out mutants were collected. The left column provided the numbers of the colonies (Col. T) of each concerned mutant. For all the colonies the loci IDs of the genes (PP) or the intergenic regions (IR) knocked out by each mini-Tn*5* derivative were given in a single row for a single mutant. The third and fifth columns described the sign allocated to the mini-Tn*5 KpF* and mini-Tn*5 TF* to inform about their relative orientation in the genome of the mutant. In the columns four and seven the detection of the Oend of each mini-Tn*5*, by AP-PCR and sequencing experiments, was reported. If the Oend was clearly found by analyzing the sequence data a 'yes' was inserted in the cell of the table. When the sequencing of the AP-PCR products provided the potential knock-out gene but no trace of the Oend, a 'no' was inserted in the cell. Finally the last column 'Putative deleted fragment' indicated the size, predicted in silico, of the deleted fragment situated between both mini-Tn*5* derivatives.

**Table 8-7: Genes and intergenic regions hit by a mini-Tn*5 TF* in *P. putida* TMT double mutants**

| Locus ID | Sign | *gene* | Product name | K/EC number | Cellular role category |
|---|---|---|---|---|---|
| **PP_0041** | (+) | *cadA-1* | heavy metal translocating P-type ATPase | K01534 Cd2+/Zn2+-exporting ATPase [EC:3.6.3.3 3.6.3.5] | **Cellular processes:** Detoxification / **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_0043** | (-) | - | CzcA family cobalt/zinc/cadmium efflux transporter permease | heavy-metal exporter, HME family | **Cellular processes:** Detoxification / **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_0065** | (-) | *trkA* | potassium transporter peripheral membrane component | trk system potassium uptake protein TrkA | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_0144** | (+) | - | insulinase family metalloprotease | - | **Protein fate:** Degradation of proteins, peptides, and glycopeptides |
| **PP_0167** | (-) | *lapB* | ATPase, ABC transporter. Secretion of LapA | ATP-binding cassette, subfamily C, bacterial LapB | **Protein fate:** Protein and peptide secretion and trafficking / **Transport and binding proteins:** Other |
| **PP_0168** | (+) | *lapA* | Large adhesion, surface associated. Biofilm formation | surface adhesion protein | **Cellular processes:** Cell adhesion |
| **PP_0179** | (+) | - | RND efflux transporter | - | **Transport and binding proteins:** Unknown substrate |
| **PP_0207** | (-) | - | nitrate ABC transporter, periplasmic nitrate-binding protein, putative | K02051 sulfonate/nitrate/taurine transport system substrate-binding protein | **Transport and binding proteins:** Anions |
| **PP_0213** | (+) | *davD* | Glutaric semialdehyde dehydrogenase | K00135 succinate-semialdehyde dehydrogenase (NADP+) [EC:1.2.1.16] | **Energy metabolism:** Amino acids and amines |
| **PP_0224** | (+) | - | DszC family monooxygenase | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_0234** | (-) | *oprE* | outer membrane porin | - | **Transport and binding proteins:** Porins |
| **PP_0235** | (+) | *lsfA* | peroxidase | - | **Cellular processes:** Detoxification / **Central intermediary metabolism:** Sulfur metabolism |
| **PP_0245** | (-) | - | S1 RNA-binding domain-containing protein | K06959 | **Unknown function:** General |
| **PP_0281** | (-) | - | polar amino acid ABC transporter inner membrane subunit | K02029 polar amino acid transport system permease protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_0287** | (+) | - | AsmA family protein | AsmA protein | **Unknown function:** General |
| **PP_0302** | (-) | - | 3-hydroxybutyryl-CoA dehydrogenase (EC:1.1.1.157) | K00074 3-hydroxybutyryl-CoA dehydrogenase [EC:1.1.1.157] | **Fatty acid and phospholipid metabolism:** Degradation |
| **PP_0323** | (-) | *soxB* | sarcosine oxidase, beta subunit family | sarcosine oxidase, subunit beta [EC:1.5.3.1] | **Energy metabolism:** Amino acids and amines |

| | | | | | |
|---|---|---|---|---|---|
| **PP_0325** | **(-)** | *sox A* | sarcosine oxidase, alpha subunit family | K00302 sarcosine oxidase, subunit alpha [EC:1.5.3.1] | **Energy metabolism:** Amino acids and amines |
| **PP_0348** | **(+)** | - | hypothetical protein | - | - |
| **PP_0350** | **(-)** | - | TonB-dependent siderophore receptor | K02014 iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_0370** | **(+)** | - | acyl-CoA dehydrogenase domain-containing protein | [EC:1.3.99.-] | **Fatty acid and phospholipid metabolism:** Degradation |
| **PP_0371** | **(+)** | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_0411** | **(-)** | - | spermidine/putrescine ABC transporter ATPas | K02052 putative spermidine/putrescine transport system ATP-binding protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_0483** | **(+)** | *uvr A* | excinuclease ABC subunit A | excinuclease ABC subunit A | **DNA metabolism:** DNA replication, recombination, and repair |
| **PP_0486** | **(-)** | - | GntR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_0489** | **(-)** | - | formate dehydrogenase, alpha subunit, authentic point mutation | - | **Energy metabolism:** Electron transport |
| **PP_0499** | **(-)** | - | conserved hypothetical protein, authentic frameshift | - | **Hypothetical proteins:** Conserved |
| **PP_0573** | **(+)** | - | hypothetical protein | - | - |
| **PP_0583** | **(+)** | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_0593** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_0597** | **(-)** | *mmsA-1* | methylmalonate-semialdehyde dehydrogenase | K00140 methylmalonate-semialdehyde dehydrogenase [EC:1.2.1.27] | **Energy metabolism:** Amino acids and amines |
| **PP_0612** | **(+)** | - | FAD dependent oxidoreductase | K03153 glycine oxidase [EC:1.4.3.19] | **Unknown function:** Enzymes of unknown specificity |
| **IR_0655-56** | **(+)** | - | fimbrial protein-related protein / polar amino acid ABC transporter inner membrane subunit | / K02029 polar amino acid transport system permease protein | **Unknown function:** General / **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_0672** | **(-)** | - | sensory box protein | - | **Regulatory functions:** Small molecule interactions |
| **PP_0716** | **(+)** | - | secretion protein HlyD family protein | - | **Protein fate:** Protein and peptide secretion and trafficking / **Transport and binding proteins:** Other |

| | | | | | |
|---|---|---|---|---|---|
| **PP_0738** | **(-)** | - | hypothetical protein | - | **Hypothetical proteins:** Domain |
| **PP_0768** | **(+)** | - | response regulator/TPR domain-containing protein | | **Signal transduction:** Two-component systems |
| **PP_0806** | **(+)** | *lap F* | seed colonization adhesion protein LapF | | **Cellular processes:** Cell adhesion |
| **PP_0825** | **(-)** | - | phosphate ABC transporter, ATP-binding protein, putative | K02041 phosphonate transport system ATP-binding protein | **Transport and binding proteins:** Anions |
| **PP_0873** | **(+)** | - | periplasmic polyamine-binding protein, putative | putrescine transport system substrate-binding | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_0887** | **(+)** | - | integral membrane sensor signal transduction histidine kinase | - | **Regulatory functions:** Protein interactions **/ Signal transduction:** Two-component systems |
| **PP_0902** | **(+)** | - | integral membrane sensor signal transduction histidine kinase | - | **Regulatory functions:** Protein interactions **/ Signal transduction:** Two-component systems |
| **PP_0913** | **(+)** | - | hypothetical protein | putative iron-regulated protein | **Hypothetical proteins:** conserved |
| **PP_0918** | **(+)** | - | NAD-dependent epimerase/dehydratase | - | **Fatty acid and phospholipid metabolism:** Other |
| **PP_0928** | **(+)** | - | Na+/Ca+ antiporter, CaCA family | K07301 | **Unknown function:** General |
| **PP_0958** | **(-)** | *ttg2 A* | toluene tolerance ABC efflux transporter, ATP-binding protein | K02065 putative ABC transport system ATP-binding protein | **Cellular processes:** Toxin production and resistance **/ Transport and binding proteins:** Unknown substrate |
| **PP_0971** | **(+)** | - | conserved hypothetical protein | Unclassified; Poorly Characterized; General function prediction only | **Hypothetical proteins:** conserved |
| **PP_0976** | **(-)** | - | putative SAM-dependent methyltransferase | K06970 ribosomal RNA large subunit methyltransferase F [EC:2.1.1.48] | **Hypothetical proteins:** Conserved |
| **PP_1018** | **(-)** | - | sugar ABC transporter, ATP-binding subunit | K02023 multiple sugar transport system ATP-binding protein | **Transport and binding proteins:** Carbohydrates, organic alcohols, and acids |
| **IR_1036-37** | **(-)** | *- / pur L* | ligand-binding/transglycosylase domain-containing protein / phosphoribosylformylglycinamidine synthase (EC:6.3.5.3) | / phosphoribosylformylglycinamidine synthase [EC:6.3.5.3] | **Unknown function:** General **/ Purines, pyrimidines, nucleosides, and nucleotides:** Purine ribonucleotide biosynthesis |
| **PP_1055** | **(+)** | *gsp N* | type II secretion pathway protein GspN | K02463 general secretion pathway protein N | **Protein fate:** Protein and peptide secretion and trafficking |
| **PP_1065** | **(-)** | - | GlpM family protein | membrane protein GlpM | **Cell envelope:** Other |
| **PP_1072** | **(-)** | - | leucine-rich repeat-containing protein | - | **Unknown function:** General |

| | | | | | |
|---|---|---|---|---|---|
| **PP_1075** | **(+)** | *glpK* | glycerol kinase | glycerol kinase [EC:2.7.1.30] | **Energy metabolism:** other |
| **PP_1076** | **(-)** | *glpF* | MIP family channel protein | K02440 glycerol uptake facilitator protein | **Transport and binding proteins:** Carbohydrates, organic alcohols, and acids |
| **PP_1090** | **(+)** | - | LuxR family DNA-binding response regulator | K07690 two-component system, NarL family, response regulator EvgA | **Regulatory functions:** DNA interactions / **Signal transduction:** Two-component systems |
| **PP_1105** | **(+)** | - | ATP-dependent DNA ligase (EC:6.5.1.1) | DNA ligase (ATP) [EC:6.5.1.1] | **DNA metabolism:** DNA replication, recombination, and repair |
| **PP_1127** | **(-)** | *estC* | beta-lactamase (carboxylesterase) | - | **Fatty acid and phospholipid metabolism:** Degradation |
| **PP_1141** | **(+)** | *braC* | extracellular ligand-binding receptor | branched-chain amino acid transport system substrate-binding protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_1154** | **(-)** | - | sensory box protein | - | **Regulatory functions:** Small molecule interactions |
| **PP_1225** | **(-)** | - | radical SAM domain protein | K10026 queuosine biosynthesis protein QueE | **Protein fate:** Protein modification and repair |
| **PP_1277** | **(-)** | *algA* | mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase | K00971 mannose-1-phosphate guanylyltransferase [EC:2.7.7.22] K01809 mannose-6-phosphate isomerase [EC:5.3.1.8] | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides |
| **PP_1281** | **(+)** | *algL* | poly(beta-D-mannuronate) lyase (EC:4.2.2.3) | poly(beta-D-mannuronate) lyase [EC:4.2.2.3] | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides |
| **PP_1283** | **(+)** | - | carbohydrate-binding and sugar hydrolysis | - | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides |
| **PP_1298** | **(-)** | *aapQ* | polar amino acid ABC transporter inner membrane subunit | K09970 general L-amino acid transport system permease protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_1308** | **(-)** | *mdeA* | methionine gamma-lyase (EC:4.4.1.11) | K01761 methionine-gamma-lyase [EC:4.4.1.11] | **Energy metabolism:** Amino acids and amines |
| **PP_1449** | **(-)** | *hlpA* | Hemolysin-like protein. Seed colonization and iron acquisition | hemolysin | **Central intermediary metabolism:** Nitrogen metabolism / **Cellular processes:** Cell adhesion |
| **PP_1492** | **(+)** | - | CheA signal transduction histidine kinase | K03407 two-component system, chemotaxis family, sensor kinase CheA [EC:2.7.13.3] | **Regulatory functions:** Small molecule interactions / **Signal transduction:** Two-component systems |
| **PP_1513** | **(-)** | - | hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_1626** | **(+)** | *mutS* | DNA mismatch repair protein MutS | K03555 DNA mismatch repair protein MutS | **DNA metabolism:** DNA replication, recombination, and repair |

| | | | | | |
|---|---|---|---|---|---|
| **PP_1645** | **(-)** | *arsC-1* | arsenate reductase | arsenate reductase [EC:1.20.4.1] | **Cellular processes:** Detoxification |
| **PP_1648** | **(-)** | - | hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_1656** | **(-)** | *relA* | (p)ppGpp synthetase I, SpoT/RelA | GTP pyrophosphokinase [EC:2.7.6.5] | **Cellular processes:** Adaptations to atypical conditions |
| **PP_1740** | **(+)** | - | beta (1-6) glucans synthase, putative | - | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides **/ Energy metabolism:** Biosynthesis and degradation of polysaccharides |
| **PP_1751** | **(-)** | *mnmC, yfcK, trmC* | 5-methylaminomethyl-2-thiouridine methyltransferase | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_1758** | **(+)** | - | hypothetical protein | K07146 | **Unknown function:** General |
| **PP_1776** | **(-)** | - | mannose-6-phosphate isomerase/mannose-1-phosphate guanylyltransferase | mannose-1-phosphate guanylyltransferase [EC:2.7.7.22] | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides **/ Energy metabolism:** Sugars |
| **PP_1816** | **(+)** | - | alcohol dehydrogenase, zinc-containing | K07119 | **Energy metabolism:** Fermentation |
| **PP_1831** | **(+)** | - | major facilitator superfamily MFS_1 | K05820 MFS transporter, PPP family, 3-phenylpropionic acid transporter | **Cell envelope:** Other |
| **PP_1838** | **(-)** | - | hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_1866** | **(+)** | - | phospho-2-dehydro-3-deoxyheptonate aldolase | 3-deoxy-7-phosphoheptulonate synthase [EC:2.5.1.54] | **Amino acid biosynthesis:** Aromatic amino acid family |
| **IR_1871-72** | **(+)** | *htpX /-* | heat shock protein HtpX / aminotransferase AlaT (EC:2.6.1.2) | K03799 heat shock protein HtpX [EC:3.4.24.-] / K10907 aminotransferase [EC:2.6.1.-] | **Protein fate:** Protein folding and stabilization **/ Unknown function:** Enzymes of unknown specificity |
| **PP_1896** | **(-)** | - | ABC transporter | antibiotic transport system permease protein | **Transport and binding proteins:** Unknown substrate |
| **PP_1952** | **(-)** | - | metallo-beta-lactamase family protein | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_1983** | **(-)** | - | sensory box protein | - | **Regulatory functions:** Small molecule interactions |
| **PP_2017** | **(+)** | *pepN* | aminopeptidase N | K01256 aminopeptidase N [EC:3.4.11.2] | **Protein fate:** Degradation of proteins, peptides, and glycopeptides |
| **PP_2032** | **(-)** | - | ATPase, putative | MoxR-like ATPase [EC:3.6.3.-] | **Unknown function:** General |

| | | | | | |
|---|---|---|---|---|---|
| **PP_2073** | **(+)** | - | acetyltransferase | K03827<br>putative acetyltransferase<br>[EC:2.3.1.-] | **Unknown function:** Enzymes of unknown specificity |
| **PP_2080** | **(+)** | - | NAD-glutamate dehydrogenase | K00260<br>glutamate dehydrogenase<br>[EC:1.4.1.2] | **Hypothetical proteins:** Conserved |
| **PP_2109** | **(+)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **IR_2118-19** | **(-)** | - | hypothetical prot/<br>lipid ABC transporter ATPase/inner membrane protein | ATP-binding cassette, subfamily B, bacterial | - |
| **PP_2144** | **(+)** | - | TetR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_2164** | **(+)** | - | type IV pilus assembly PilZ | - | **Cell envelope:** Surface structures |
| **PP_2179** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** conserved |
| **PP_2180** | **(-)** | - | putative aminotransferase | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_2209** | **(+)** | *phnW* | 2-aminoethylphosphonate--pyruvate transaminase | K03430<br>2-aminoethylphosphonate-pyruvate transaminase<br>[EC:2.6.1.37] | **Energy metabolism:** Other |
| **PP_2270** | **(-)** | - | DNA primase/helicase | - | **Mobile and extrachromosomal element functions:** Prophage functions |
| **PP_2271** | **(+)** | - | hypothetical protein | - | - |
| **PP_2281** | **(-)** | - | capsid assembly protein, putative | - | **Mobile and extrachromosomal element functions:** Prophage functions |
| **PP_2286** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_2334** | **(+)** | *prpB* | 2-methylisocitrate lyase (EC:4.1.3.30) | K03417<br>methylisocitrate lyase<br>[EC:4.1.3.30] | **Energy metabolism:** Aerobic |
| **PP_2340** | **(-)** | - | conserved hypothetical protein | K06938 | **Hypothetical proteins:** Conserved |
| **IR_2348-49** | **(+)** | - | integral membrane sensor signal transduction histidine kinase / CBS domain-containing protein | / K07168<br>CBS domain-containing membrane protein | **[Regulatory functions:** Protein interactions **/ Signal transduction:** Two-component systems**] /**<br>**Unknown function:** General |
| **PP_2356** | **(+)** | - | phytochrome family protein, putative | - | **Cellular processes:** Adaptations to atypical conditions |
| **PP_2369** | **(-)** | - | leucine-rich repeat-containing protein | - | **Unknown function:** General |

| | | | | | |
|---|---|---|---|---|---|
| **PP_2395** | **(+)** | - | leucine-rich repeat-containing protein | - | **Unknown function:** General |
| **PP_2404** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** conserved |
| **PP_2409** | **(+)** | - | cobalt-zinc-cadmium resistance protein CzcB, putative | - | **Cellular processes:** Detoxification / **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_2428** | **(+)** | - | sugar efflux transporter | MFS transporter, DHA1 family, purine ribonucleoside efflux pump | **Transport and binding proteins:** Unknown substrat**e** |
| **PP_2444** | **(+)** | - | LysR family transcriptional regulator | K03717 LysR family transcriptional regulator, transcriptional activator of nhaA | **Regulatory functions:** DNA interactions |
| **PP_2453** | **(+)** | *ans A* | L-asparaginase, type II | K05597 glutamin-(asparagin-)ase [EC:3.5.1.38] | **Energy metabolism:** Amino acids and amines |
| **PP_2534** | **(+)** | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **IR_2541-42** | **(+)** | - | transcriptional factor-related protein / GntR family transcriptional regulator | - | **[Unknown function:** General / **Transport and binding proteins:** Amino acids, peptides and amines**] / Regulatory functions:** DNA interactions |
| **PP_2542** | **(-)** | - | GntR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_2554** | **(+)** | - | 4-hydroxyphenylpyruvate dioxygenase | K00457 4-hydroxyphenylpyruvate dioxygenase [EC:1.13.11.27] | **Energy metabolism: Amino acids and amines** |
| **PP_2558** | **(+)** | - | TolC family type I secretion outer membrane protein | outer membrane protein HasF | **Protein fate:** Protein and peptide secretion and trafficking |
| **PP_2561** | **(-)** | - | heme peroxidase | - | **Cellular processes:** Toxin production and resistance **/ Protein fate:** Protein and peptide secretion and trafficking |
| **PP_2576** | **(-)** | - | PqiA family protein, authentic frameshift | - | **Unknown function:** General |
| **PP_2588** | **(-)** | - | aminotransferase, class III (putative aminotransferase) | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_2590** | **(-)** | - | outer membrane ferric siderophore receptor | K02014 iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_2600** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_2613** | **(-)** | - | hypothetical protein | - | - |
| **PP_2642** | **(-)** | - | GntR family transcriptional regulator | K00375 GntR family transcriptional regulator / MocR family aminotransferase | **Regulatory functions:** DNA interactions |

| | | | | | |
|---|---|---|---|---|---|
| **PP_2645** | **(+)** | *mgt B* | magnesium-translocating P-type ATPase | K01531 Mg2+-importing ATPase [EC:3.6.3.2] | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_2650** | **(+)** | *gbd* | iron-containing alcohol dehydrogenase | K00100 [EC:1.1.1.-] | **Energy metabolism:** Other |
| **PP_2677** | **(-)** | - | hypothetical protein | - | - |
| **PP_2695** | **(-)** | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_2696** | **(-)** | *met R-2* | transcriptional activator MetR | - | **Regulatory functions:** DNA interactions |
| **IR_2702-03** | **(+)** | - | hypothetical protein / major facilitator transporter | - | **Transport and binding proteins:** Porins **/ Transport and binding proteins:** Carbohydrates, organic alcohols, and acids |
| **PP_2703** | **(+)** | - | major facilitator transporter | - | **Transport and binding proteins:** Carbohydrates, organic alcohols, and acids |
| **PP_2717** | **(-)** | *Ars B-2* | arsenite efflux transporter | K03893 arsenical pump membrane protein | **Cellular processes:** Detoxification/ **Transport and binding proteins:** Others |
| **PP_2749** | **(-)** | - | branched-chain amino acid ABC transporter, permease protein, putative | K01997 branched-chain amino acid transport system permease protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_2762** | **(+)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_2787** | **(+)** | - | transporter, putative | K07003 | **Transport and binding proteins:** Unknown substrate |
| **PP_2812** | **(+)** | - | transporter, putative | K07003 | **Transport and binding proteins:** Unknown substrate |
| **PP_2822** | **(+)** | - | hypothetical protein | - | - |
| **PP_2841** | **(+)** | - | phage integrase family site specific recombinase | - | **DNA metabolism:** DNA replication, recombination, and repair **/ Mobile and extrachromosomal element functions:** Prophage functions |
| **PP_2853** | **(+)** | | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_2859** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_2870** | **(+)** | - | PotD/PotF family extracellular solute-binding protein | putative spermidine/putrescine transport system substrate-binding protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_2900** | **(-)** | - | hypothetical protein | - | - |
| **PP_2914** | **(-)** | *pro P* | proline/glycine betaine transporter | K03762 MFS transporter, MHS family, proline/betaine | **Transport and binding proteins:** Amino acids, peptides and amines |

| | | | | | |
|---|---|---|---|---|---|
| | | | | transporter | |
| **PP_2991** | **(-)** | - | hypothetical protein | - | **-** |
| **IR_2994-45** | **(-)** | - | (NADH:flavin oxidoreductase/NADH oxidase) / DNA topology modulation kinase FlaR, putative | - | **Unknown function:** Enzymes of unknown specificity **/ [Cellular processes:** Chemotaxis and motility**/Regulatory functions:** DNA interactions**]** |
| **PP_3038** | **(-)** | - | TraC domain-containing protein | - | **Cellular processes:** Toxin production and resistance **/ Mobile and extrachromosomal element functions:** Prophage functions |
| **PP_3074** | **(+)** | - | membrane protein putative | - | **Cell envelope:** Other |
| **PP_3124** | **(-)** | - | short-chain fatty acid transporter family protein | K02106 short-chain fatty acids transporter | **Transport and binding proteins:** Carbohydrates, organic alcohols, and acids |
| **IR_3152-53** | **(-)** | - | hypothetical protein / transcriptional regulator, LysR family | - | **Regulatory functions:** DNA interactions **/** |
| **PP_3224** | **(+)** | - | aldolase II superfamily protein | - | **Unknown function:** General |
| **IR_3232-33** | **(+)** | - | acetyltransferase / Crp/FNR family transcriptional regulator | - | **Unknown function:** Enzymes of unknown specificity **/ Regulatory functions:** DNA interactions |
| **PP_3251** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_3260** | **(-)** | *lig D* | ATP-dependent DNA ligase (EC:6.5.1.1) | DNA ligase (ATP) [EC:6.5.1.1] | **DNA metabolism:** DNA replication, recombination, and repair |
| **IR_3260-3261** | **(-)** | *lig D* / - | ATP-dependent DNA ligase (EC:6.5.1.1) / hypothetical protein | DNA ligase (ATP) [EC:6.5.1.1] / | **DNA metabolism:** DNA replication, recombination, and repair **/ Hypothetical proteins:** Conserved |
| **PP_3262** | **(+)** | - | metallothionein, putative | - | **Cellular processes:** Detoxification **/Transport and binding proteins:** Cations and iron carrying compounds |
| **IR_3273-74** | **(+)** | -/ *pha I* | conserved hypothetical protein / phenylacetate-CoA oxygenase/reductase, PaaK subunit | / K02613 phenylacetic acid degradation NADH oxidoreductase | **Hypothetical proteins:** Conserved / **Energy metabolism**: Other |
| **PP_3279** | **(+)** | *pha E* | phenylacetate-CoA ligase | phenylacetate-CoA ligase [EC:6.2.1.30] | **Energy metabolism:** Other |
| **PP_3315** | **(+)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_3327** | **(+)** | - | proton/sodium-glutamate/aspartate symporter, authentic frameshift | - | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_3329** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |

| | | | | | |
|---|---|---|---|---|---|
| **PP_3348** | **(+)** | - | GGDEF domain-containing protein | - | **Unknown function:** general |
| **PP_3364** | **(-)** | - | response regulator | - | **Regulatory functions:** Protein interactions **/ Signal transduction:** Two-component systems |
| **PP_3396** | **(+)** | - | hypothetical protein | - | - |
| **PP_3439** | **(+)** | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_3450** | **(-)** | - | hypothetical protein | bacteriophage N4 adsorption protein A | - |
| **PP_3483** | **(+)** | - | type II secretion system protein E | K02454 general secretion pathway protein E | **Protein fate**: Protein and peptide secretion and trafficking |
| **PP_3490** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_3498** | **(-)** | - | ISPpu11, transposase | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| **PP_3514** | **(+)** | *hyuB* | hydantoinase B/oxoprolinase | K01474 N-methylhydantoinase B [EC:3.5.2.14] | **Amino acid biosynthesis:** Other |
| **PP_3529** | **(+)** | - | monooxygenase, putative | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_3539** | **(-)** | - | transcriptional regulator, putative | - | **Regulatory functions:** DNA interactions |
| **PP_3570** | **(+)** | - | carbohydrate-selective porin OprB | - | **Transport and binding proteins:** Porins |
| **PP_3578** | **(-)** | - | phosphoglucomutase, alpha-D-glucose phosphate-specific | phosphoglucomutase [EC:5.4.2.2] | **Energy metabolism:** sugars |
| **PP_3583** | **(-)** | - | RND efflux transporter, authentic frameshift | - | **Transport and binding proteins:** Unknown substrate |
| **PP_3589** | **(-)** | *sdaC* | aromatic amino acid ABC transporter permease | serine transporter | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_3597** | **(-)** | - | amino acid ABC transporter ATP-binding protein | K02028 polar amino acid transport system ATP-binding protein [EC:3.6.3.21] | **Transport and binding proteins:** Amino acids, peptides and amines |
| **IR_3603-04** | **(-)** | - | GntR family transcriptional regulator / hypothetical protein | K05799 GntR family transcriptional regulator, transcriptional repressor for pyruvate dehydrogenase complex / | **Regulatory functions:** DNA interactions **/** |
| **PP_3612** | **(-)** | - | TonB-dependent siderophore receptor | K02014 iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |

| | | | | | |
|---|---|---|---|---|---|
| **PP_3625** | **(+)** | - | amino acid transporter LysE | - | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_3660** | **(-)** | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_3733** | **(+)** | - | ABC transporter lipoprotein, putative | - | **Transport and binding proteins:** Unknown substrate |
| **PP_3786** | **(-)** | - | aminotransferase | - | **Cellular processes:** Toxin production and resistance |
| **IR_3795-96** | **(+)** | - | hypothetical protein / L-ornithine N5-oxygenase | / L-ornithine N5-oxygenase [EC:1.13.12.-] | **/ Central intermediary metabolism:** Other |
| **PP_3796** | **(+)** | - | L-ornithine N5-oxygenase | L-ornithine N5-oxygenase [EC:1.13.12.-] | **Central intermediary metabolism:** Other |
| **PP_3827** | **(-)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_3846** | **(-)** | - | carbon-nitrogen hydrolase family protein | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_3869** | **(+)** | - | phage sheath protein, putative | - | **Mobile and extrachromosomal element functions:** Prophage functions |
| **PP_3900** | **(-)** | *hicA-2* | HicA protein | - | **Unknown function:** General |
| **PP_3915** | **(+)** | - | hypothetical protein | - | - |
| **PP_3970** | **(-)** | - | alcohol dehydrogenase GroES domain protein | - | **Energy metabolism:** Fermentation |
| **PP_3988** | **(+)** | - | hypothetical protein | - | - |
| **PP_4011** | **(+)** | *icd* | isocitrate dehydrogenase, NADP-dependent | K00031 isocitrate dehydrogenase [EC:1.1.1.42] | **Energy metabolism:** TCA cycle |
| **PP_4019** | **(+)** | *topB* | DNA topoisomerase III | K03169 DNA topoisomerase III [EC:5.99.1.2] | **DNA metabolism:** DNA replication, recombination, and repair |
| **PP_4029** | **(-)** | *nudC* | NADH pyrophosphatase | K03426 NAD+ diphosphatase [EC:3.6.1.22] | **Central intermediary metabolism:** Other |
| **PP_4031** | **(-)** | *nhaB* | sodium/proton antiporter | Na+:H+ antiporter, NhaB family | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_4060** | **(-)** | - | alpha-amylase family protein | - | **Energy metabolism:** Biosynthesis and degradation of polysaccharides |
| **PP_4075** | **(+)** | - | conserved hypothetical protein, authentic point mutation | - | **Hypothetical proteins:** Conserved |

| | | | | | |
|---|---|---|---|---|---|
| **PP_4111** | **(+)** | *fus A* | elongation factor G | elongation factor EF-G [EC:3.6.5.3] | **Protein synthesis:** Translation factors |
| **PP_4146** | **(+)** | - | peptide ABC transporter, periplasmic peptide-binding protein | peptide/nickel transport system substrate-binding protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_4217** | **(-)** | *fpv A* | outer membrane ferripyoverdine receptor | iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_4220** | **(+)** | - | non-ribosomal peptide synthetase | - | **Unknown function:** General |
| **PP_4223** | **(-)** | - | diaminobutyrate--2-oxoglutarate aminotransferase (EC:2.6.1.76) | diaminobutyrate-2-oxoglutarate transaminase [EC:2.6.1.76] | **Central intermediary metabolism:** Other |
| **PP_4224** | **(-)** | - | integral membrane sensor signal transduction histidine kinase | two-component system, OmpR family, sensor kinase [EC:2.7.13.3] | **Regulatory functions:** Protein interactions / **Signal transduction:** Two-component systems |
| **PP_4278** | **(+)** | *xdh A* | xanthine dehydrogenase, XdhA subunit | xanthine dehydrogenase [EC:1.17.1.4] | **Purines, pyrimidines, nucleosides, and nucleotides:** Salvage of nucleosides and nucleotides / **Energy metabolism:** Other |
| **PP_4316** | **(-)** | - | 2-hydroxyacid dehydrogenase | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_4330** | **(-)** | - | toxin-related protein | - | **Unknown function:** General |
| **PP_4358** | **(-)** | *fli M* | flagellar motor switch protein FliM | K02416 flagellar motor switch protein FliM | **Cellular processes:** Chemotaxis and motility |
| **PP_4379** | **(+)** | - | beta-ketoacyl-acyl-carrier-protein synthase I | K00648 3-oxoacyl-[acyl-carrier-protein] synthase III [EC:2.3.1.180] | **Fatty acid and phospholipid metabolism:** Biosynthesis |
| **PP_4441** | **(-)** | - | ISPpu14, transposase Orf1 | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| **IR_4446-47** | **(+)** | - | group II intron-encoding maturase, putative / hypothetical protein | - | **Mobile and extrachromosomal element functions:** Transposon functions / |
| **PP_4455** | **(-)** | - | opine ABC transporter, periplasmic binding protein, putative | K02033 peptide/nickel transport system permease protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| **IR_4467-68** | **(-)** | - | LysR family transcriptional regulator / Cro/CI family transcriptional regulator | - | **Regulatory functions:** DNA interactions / **Regulatory functions:** DNA interactions |
| **PP_4468** | **(+)** | - | Cro/CI family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_4470** | **(-)** | *alg Z* | Arc domain protein DNA binding domain protein(alginate biosynthesis transcriptional activator) | - | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides / **Regulatory functions:** DNA interactions |
| **PP_4475** | **(-)** | *ast E* | succinylglutamate desuccinylase | K05526 succinylglutamate desuccinylase [EC:3.5.1.96] | **Energy metabolism:** Amino acids and amines |

| | | | | | |
|---|---|---|---|---|---|
| **PP_4487** | **(-)** | *acs A* | acetyl-CoA synthetase (EC:6.2.1.1) | K01895 acetyl-CoA synthetase [EC:6.2.1.1] | **Energy metabolism:** (Fermentation / Other) |
| **PP_4516** | **(+)** | *rec Q* | ATP-dependent DNA helicase RecQ | K03654 ATP-dependent DNA helicase RecQ [EC:3.6.1.-] | **DNA metabolism:** DNA replication, recombination, and repair |
| **PP_4610** | **(-)** | - | PepSY-associated TM helix domain protein | - | **Cell envelope:** Other |
| **PP_4626** | **(+)** | - | LrgA family protein | K06518 holin-like protein | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides |
| **PP_4644** | **(+)** | *rad A* | DNA repair protein RadA | K04485 DNA repair protein RadA/Sms | **DNA metabolism:** DNA replication, recombination, and repair |
| **PP_4683** | **(+)** | *mrc B* | penicillin-binding protein 1B | K05365 penicillin-binding protein 1B [EC:2.4.1.129 3.4.-.-] | **Cell envelope:** Biosynthesis and degradation of murein sacculus and peptidoglycan |
| **PP_4688** | **(-)** | - | transport system permease protein | iron complex transport system permease protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_4736** | **(-)** | *lld D* | L-lactate dehydrogenase (EC:1.1.2.3) | K00101 L-lactate dehydrogenase (cytochrome) [EC:1.1.2.3] | **Energy metabolism:** Fermentation / **Energy metabolism:** Glycolysis/gluconeogenesis |
| **IR_4771-72** | **(+)** | *- / hrp B* | hypothetical protein / ATP-dependent helicase HrpB | / K03579 ATP-dependent helicase HrpB [EC:3.6.1.-] | **Cell envelope:** Other / **Transcription:** Other |
| **PP_4830** | **(+)** | *cob L* | precorrin-6y C5,15-methyltransferase (decarboxylating), CbiE subunit | precorrin-6Y C5,15-methyltransferase / precorrin-8W decarboxylase [EC:2.1.1.132 1.-.-.-] | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Heme, porphyrin, and cobalamin |
| **PP_4852** | **(-)** | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **PP_4916** | **(-)** | - | conserved hypothetical protein | K07000 | **Hypothetical proteins:** Conserved |
| **PP_4918** | **(+)** | - | conserved hypothetical protein | K09920 hypothetical protein | **Hypothetical proteins:** Conserved |
| **PP_4946** | **(-)** | *put P* | sodium/proline symporter | sodium/proline symporter | **Transport and binding proteins:** Amino acids, peptides and amines |
| **PP_4954** | **(+)** | - | integral membrane protein, YccS/YhfK family | - | **Cell envelope:** Other |
| **IR_4979-80** | **(+)** | - | periplasmic binding protein, putative / DEAD-box ATP dependent DNA helicase | K02030 polar amino acid transport system substrate-binding protein / K11927 ATP-dependent RNA helicase RhlE [EC:3.6.1.-] | **Transport and binding proteins:** Unknown substrate / **Transcription:** Other |
| **PP_5009** | **(+)** | - | hypothetical protein | - | - |
| **PP_5026** | **(+)** | *mdoG, opgG* | glucan biosynthesis protein G | periplasmic glucans biosynthesis protein | **Cell envelope:** Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides |

| | | | | | |
|---|---|---|---|---|---|
| **PP_5045** | **(+)** | *thiI* | thiamine biosynthesis protein ThiI | K03151 thiamine biosynthesis protein ThiI | **Biosynthesis of cofactors,** prosthetic groups, and carriers: Thiamine |
| **PP_5067** | **(-)** | - | potassium efflux protein KefA | potassium efflux system protein KefA | **Transport and binding proteins:** Cations and iron carrying compounds |
| **PP_5068** | **(+)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_5091** | **(+)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **PP_5093** | **(-)** | *pil T* | twitching motility protein | twitching motility protein | **Cell envelope:** Surface structures / **Cellular processes:** (Chemotaxis and motility / Cell adhesion) |
| **PP_5154** | **(-)** | - | FAD linked oxidase domain protein | - | **Unknown function:** Enzymes of unknown specificity |
| **PP_5165** | **(+)** | - | NlpA lipoprotein | K02073 D-methionine transport system substrate-binding protein | **Cell envelope:** Other |
| **IR_5172-73** | **(-)** | - | hypothetical protein / acriflavin resistance protein | - | **Hypothetical proteins:** Conserved / **[Cellular processes:** Toxin production and resistance / **Transport and binding proteins:** Unknown substrate**]** |
| **PP_5188** | **(+)** | - | transcriptional regulator BkdR, putative | K05800 Lrp/AsnC family transcriptional regulator | **Regulatory functions:** DNA interactions |
| **PP_5192** | **(-)** | *gcv P-2* | glycine dehydrogenase (EC:1.4.4.2) | K00282 glycine dehydrogenase subunit 1 [EC:1.4.4.2] K00283 glycine dehydrogenase subunit 2 [EC:1.4.4.2] | **Energy metabolism:** Amino acids and amines |
| **PP_5241** | **(+)** | - | LuxR family DNA-binding response regulator | - | **Regulatory functions:** DNA interactions **/ Signal transduction:** Two-component systems |
| **PP_5252** | **(+)** | - | amidohydrolase 3 | K07047 | **Central intermediary metabolism:** Other / **Unknown function:** Enzymes of unknown specificity |
| **PP_5261** | **(-)** | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| **IR_5275-76** | **(-)** | - | GntR family transcriptional regulator / (phospholipase D/transphosphatidylase) | - | **[Regulatory functions:** DNA interactions **/ Unknown function:** Enzymes of unknown specificity**] / Unknown function:** general |
| **PP_5318** | **(+)** | *ubi A* | 4-hydroxybenzoate octaprenyltransferase | K03179 4-hydroxybenzoate octaprenyltransferase [EC:2.5.1.-] | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Menaquinone and ubiquinone |
| **PP_5334** | **(+)** | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| **IR_5339-40** | **(-)** | - | AraC family transcriptional regulator/ histone deacetylase superfamily | - | **Regulatory functions:** DNA interactions **/ Energy metabolism:** Amino acids and amines |
| **IR_5347-48** | **(+)** | *acc C-2 /-* | acetyl-CoA carboxylase subunit A (EC:6.4.1.2) / LysR family | pyruvate carboxylase subunit A [EC:6.4.1.1] / | **Fatty acid and phospholipid metabolism:** Biosynthesis / **Regulatory functions:** DNA |

| | | | | | |
|---|---|---|---|---|---|
| | | | transcriptional regulator | | interactions |
| **PP_5366** | **(+)** | *lpd 3* | dihydrolipoamide dehydrogenase (EC:1.8.1.4) | dihydrolipoamide dehydrogenase [EC:1.8.1.4] | **Energy metabolism:** ( Pyruvate dehydrogenase/TCA cycle) |
| **PP_5372** | **(+)** | - | aldehyde dehydrogenase family protein | - | **Energy metabolism:** Fermentation |
| **PP_5405** | **(+)** | - | transposase, TnsB-related protein | - | **Mobile and extrachromosomal element functions:** Transposon functions |

After insertion by triparental mating of the mini-Tn*5 TF* in the pool of nine *P. putida* SMT strains, 573 exconjugants were selected for their {$Km^R$ and $Tel^R$} phenotype. Out of these 573 strains 278 were submitted to AP-PCR amplification and sequencing in order to describe the knocked out genes and intergenic regions. In total 252 independent regions were hit by a mini-Tn*5 TF* in the different strains. The list of genes and intergenic regions was established and reported in the Table 8-7. The first column indicated which gene or intergenic region was concerned by the analysis. In the second column the orientation of the mini-Tn*5 TF* inside the ORFs was given, symbolized either by a (+) or a (-) sign in function of the sequence found directly downstream the Oend of the mini-transposon. The four next columns provided information about the gene itself and the encoded protein. The name of the gene and the name of the product were found in the 3rd and 4th column, respectively. K and/or EC numbers related to the protein were annotated in the 5th column. Finally the cellular role category corresponding in which the encoded protein played a role was reported in the last column. When an intergenic region was extracted from the sequencing results, the data related to the upstream and downstream genes were related in a single row. The colors used in the table were related to the different cellular role categories listed hereafter:

| | |
|---|---|
| 2 | Amino acid Biosynthesis |
| 3 | Biosynthesis of cofactors, prosthetic groups and carriers |
| 17 | Cell envelope |
| 18 | Cellular processes |
| 7 | Central intermediary metabolism |
| 8 | DNA metabolism |
| 26 | Energy metabolism |
| 5 | Fatty acid and phospholipid metabolism |
| 8 | Mobile and extrachromosomal element functions |
| 9 | Protein fate |
| 1 | Protein synthesis |

| | |
|---|---|
| 1 | Purines, pyrimidines, nucleosides, and nucleotides |
| 28 | Regulatory functions |
| 8 | Signal transduction |
| 47 | Transport and binding proteins |

The genes encoding proteins with unknown functions or hypothetical proteins were left in white. In the same way the proteins participating to more than one cellular role category were not colored if the categories were different. The numbers inscribed in the colored rectangles indicated the number of proteins participating in the corresponding cellular role categories.

## 8.4 List of knockout genes in the different *P. putida* $\Delta_1$ mutants

**Table 8-8: List of knockout genes in the *P. putida* $\Delta_1$-91 mutant**

| Locus ID | *gene* | Product name | K/EC numbers | **Cellular role category** |
|---|---|---|---|---|
| PP_3490 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3491 | - | enoyl-coenzyme A hydratase/isomerase family protein | enoyl-CoA hydratase [EC:4.2.1.17] | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3492 | *acdA* | acyl-CoA dehydrogenase domain-containing protein | [EC:1.3.99.-] | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3493 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3494 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3495 | - | conserved hypothetical protein | K06967 | **Hypothetical proteins:** Conserved |
| PP_3496 | - | endoribonuclease L-PSP | - | **Transcription:** Degradation of RNA |
| PP_3497 | - | U32 family peptidase | putative protease [EC:3.4.-.-] | **Protein fate:** Degradation of proteins, peptides, and glycopeptides |
| PP_3498 | - | ISPpu11, transposase | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| PP_3499 | - | ISPpu14, transposase Orf1 | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| PP_3500 | - | ISPpu14, transposase Orf2 | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| PP_3501 | - | ISPpu14, transposase Orf3 | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| PP_3502 | - | ISPpu10, transposase | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| PP_3503 | - | sigma54 specific transcriptional regulator, fis family | - | **Regulatory functions:** Protein interactions |

| | | | | |
|---|---|---|---|---|
| PP_3504 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3505 | - | conserved hypothetical protein | magnesium chelatase subunit D [EC:6.6.1.1] | **Hypothetical proteins:** Conserved |
| PP_3506 | - | magnesium chelatase | magnesium chelatase subunit I [EC:6.6.1.1] | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Heme, porphyrin, and cobalamin **/Transport and binding proteins:** Cations and iron carrying compounds |
| PP_3507 | *cobN* | cobaltochelatase subunit CobN (EC:6.6.1.2) | cobaltochelatase CobN [EC:6.6.1.2] | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Heme, porphyrin, and cobalamin |
| PP_3508 | *cobW* | cobalamin biosynthesis protein CobW | cobalamin biosynthesis protein CobW | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Heme, porphyrin, and cobalamin |
| PP_3509 | - | glyoxalase family protein | - | **Unknown function:** Enzymes of unknown specificity |
| PP_3510 | - | hypothetical protein | - | - |
| PP_3511 | *ilvE* | branched-chain amino acid aminotransferase (EC:2.6.1.42) | branched-chain amino acid aminotransferase [EC:2.6.1.42] | **Amino acid biosynthesis:** Pyruvate family |
| PP_3512 | - | transmembrane pair domain protein | - | **Cell envelope: Other** |
| PP_3513 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3514 | *hyuB* | hydantoinase B/oxoprolinase | N-methylhydantoinase B [EC:3.5.2.14] | **Amino acid biosynthesis:** Other |
| PP_3515 | *hyuA* | 5-oxoprolinase (ATP-hydrolyzing) | N-methylhydantoinase A [EC:3.5.2.14] | **Amino acid biosynthesis:** Other |
| PP_3516 | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3517 | - | hypothetical protein | - | - |
| PP_3518 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3519 | - | putative lipoprotein | - | **Cell envelope:** Other |
| PP_3520 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3521 | - | membrane protein, putative | - | **Cell envelope:** Other |
| PP_3522 | - | endoribonuclease L-PSP | - | **Transcription:** Degradation of RNA |
| PP_3523 | - | hypothetical protein | - | - |
| PP_3524 | - | hypothetical protein | - | - |
| PP_3525 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |

| | | | | |
|---|---|---|---|---|
| PP_3526 | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3527 | - | TetR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3528 | - | aliphatic sulfonate ABC transporter periplasmic ligand-binding protein | sulfonate/nitrate/taurine transport system substrate-binding protein | **Transport and binding proteins:** Unknown substrate |
| PP_3529 | - | monooxygenase, putative | - | **Unknown function:** Enzymes of unknown specificity |

**Table 8-9: List of knockout genes in the *P. putida* $\Delta_1$-407 mutant**

| Locus ID | *gene* | Product name | K/EC numbers | **Cellular role category** |
|---|---|---|---|---|
| PP_3534 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3535 | *ggt-1* | gamma-glutamyltransferase | K00681 gamma-glutamyltranspeptidase [EC:2.3.2.2] | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Glutathione and analogs |
| PP_3536 | - | hypothetical protein | - | - |
| PP_3537 | *pobA* | 4-hydroxybenzoate 3-monooxygenase (EC:1.14.13.2) | K00481 p-hydroxybenzoate 3-monooxygenase [EC:1.14.13.2] | **Energy metabolism:** Other |
| PP_3538 | *pobR* | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3539 | - | transcriptional regulator, putative | - | **Regulatory functions:** DNA interactions |
| PP_3540 | *mvaB* | hydroxymethylglutaryl-CoA lyase (EC:4.1.3.4) | K01640 hydroxymethylglutaryl-CoA lyase [EC:4.1.3.4] | **Central intermediary metabolism:** Other |
| PP_3541 | - | Mg2+ transporter | K07507 putative Mg2+ transporter-C (MgtC) family protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| PP_3542 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3543 | - | iron-sulfur cluster-binding protein | - | **Energy metabolism:** Electron transport |
| PP_3544 | - | GntR family transcriptional regulator | - | **Regulatory functions:** DNA interactions / **Unknown function:** Enzymes of unknown specificity |
| PP_3545 | - | PAS/PAC sensor hybrid histidine kinase | - | **Regulatory functions:** Small molecule interactions / **Signal transduction:** Two-component systems |
| PP_3546 | - | PAS/PAC sensor hybrid histidine kinase | - | **Regulatory functions:** Small molecule interactions / **Signal transduction:** Two-component systems |
| PP_3547 | - | short chain dehydrogenase/reductase family oxidoreductase | - | **Unknown function:** Enzymes of unknown specificity |
| PP_3548 | - | EmrB/QacA family drug resistance transporter | K03446 MFS transporter, DHA2 family, multidrug resistance | **Cellular processes:** Toxin production and resistance / **Transport and binding proteins:** Other |

| | | | | |
|---|---|---|---|---|
| | | | protein B | |
| PP_3549 | *emrA* | secretion protein HlyD family protein | K03543 multidrug resistance protein A | **Cellular processes:** Toxin production and resistance / **Transport and binding proteins:** Other |
| PP_3550 | - | MarR family transcriptional regulator | K03712 MarR family transcriptional regulator | **Cellular processes:** Toxin production and resistance / **Regulatory functions:** DNA interactions |
| PP_3551 | - | LuxR family DNA-binding response regulator | - | **Regulatory functions**: DNA interactions / **Signal transduction:** Two-component systems |
| PP_3552 | - | PAS/PAC sensor signal transduction histidine kinase | - | **Regulatory functions:** Protein interactions / **Signal transduction:** Two-component systems |
| PP_3553 | - | acyl-CoA synthetase (EC:2.3.1.86) | K00666 fatty-acyl-CoA synthase [EC:6.2.1.-] | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3554 | - | acyl-CoA dehydrogenase domain-containing protein | - | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3555 | - | anti-FecI sigma factor, FecR | - | **Regulatory functions:** Other |
| PP_3556 | - | Na+/H+ antiporter NhaC | K07084 | **Transport and binding proteins:** Cations and iron carrying compounds |
| PP_3557 | - | methyl-accepting chemotaxis transducer | K03406 methyl-accepting chemotaxis protein | **Cellular processes:** Chemotaxis and motility |
| PP_3558 | - | substrate-binding region of ABC-type glycine betaine transport system | K02002 glycine betaine/proline transport system substrate-binding protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3559 | - | binding-protein-dependent transport systems inner membrane component | K02001 glycine betaine/proline transport system permease protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3560 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3561 | - | auxin efflux carrier | K07088 | **Fatty acid and phospholipid metabolism:** Biosynthesis |
| PP_3562 | - | hypothetical protein | - | - |
| PP_3563 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3564 | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3565 | - | amino acid transporter LysE | - | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3566 | - | major facilitator family transporter | - | **Transport and binding proteins:** Unknown substrate |
| PP_3567 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3568 | - | conserved hypothetical protein | K11312 cupin 2 domain-containing protein | **Hypothetical proteins:** Conserved |

| | | | | |
|---|---|---|---|---|
| PP_3569 | - | quinate dehydrogenase (pyrroloquinoline-quinone), putative | K05358 quinate dehydrogenase (pyrroloquinoline-quinone) [EC:1.1.99.25] | **Energy metabolism:** Other |
| PP_3570 | - | carbohydrate-selective porin OprB | - | **Transport and binding proteins:** Porins |
| PP_3571 | - | acetylornithine deacetylase (EC:3.5.1.16) | K01438 acetylornithine deacetylase [EC:3.5.1.16] | **Amino acid biosynthesis:** Glutamate family |
| PP_3572 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3573 | - | monooxygenase, putative | K07222 putative flavoprotein involved in K+ transport | **Unknown function:** Enzymes of unknown specificity |
| PP_3574 | - | endoribonuclease L-PSP | - | **Transcription:** Degradation of RNA |
| PP_3575 | - | TonB-dependent siderophore receptor | K02014 iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| PP_3576 | - | anti-FecI sigma factor, FecR | - | **Regulatory functions:** Other |
| PP_3577 | - | ECF subfamily RNA polymerase sigma factor | K03088 RNA polymerase sigma-70 factor, ECF subfamily | **Transcription:** Transcription factors |
| PP_3578 | - | phosphoglucomutase (EC:5.4.2.2) | K01835 phosphoglucomutase [EC:5.4.2.2] | **Energy metabolism:** Sugars |
| PP_3579 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3580 | - | hypothetical protein | - | - |
| PP_3581 | - | GGDEF domain-containing protein | - | **Unknown function:** General |
| PP_3582 | - | RND efflux transporter | - | **Transport and binding proteins:** Unknown substrate |
| PP_3583 | - | acriflavin resistance protein | K07789 RND superfamily, multidrug transport protein MdtC | **Transport and binding proteins:** Unknown substrate |
| PP_3584 | - | RND efflux transporter, authentic frameshift | - | **Cellular processes:** Toxin production and resistance / **Transport and binding proteins:** Other |
| PP_3585 | - | efflux transporter, RND family, MFP subunit | K07799 putative multidrug efflux transporter MdtA | **Cellular processes:** Toxin production and resistance / **Transport and binding proteins:** Other |
| PP_3586 | - | ISPpu9, transposase | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| PP_3587 | *tpx* | redoxin domain protein | K11065 thiol peroxidase, atypical 2-Cys peroxiredoxin [EC:1.11.1.15] | **Cellular processes:** Detoxification |
| PP_3588 | - | Bcr/CflA family multidrug resistance transporter | - | **Cellular processes:** Toxin production and resistance / **Transport and binding proteins:** Other |
| PP_3589 | *sdaC* | aromatic amino acid ABC transporter permease | K03837 serine transporter | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3590 | *tyrB-2* | aromatic amino acid aminotransferase (EC:2.6.1.57) | K00832 aromatic-amino-acid | **Amino acid biosynthesis:** Aromatic amino acid family |

| | | | | |
|---|---|---|---|---|
| | | | transaminase [EC:2.6.1.57] | |
| PP_3591 | - | malate/L-lactate dehydrogenase | K00025 malate dehydrogenase [EC:1.1.1.37] | **Energy metabolism:** Amino acids and amines |
| PP_3592 | - | RpiR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3593 | - | amino acid ABC transporter, periplasmic amino acid-binding protein | K02030 polar amino acid transport system substrate-binding protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3594 | - | polar amino acid ABC transporter inner membrane subunit | K02029 polar amino acid transport system permease protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3595 | - | polar amino acid ABC transporter inner membrane subunit | K02029 polar amino acid transport system permease protein | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3596 | - | FAD dependent oxidoreductase | K00285 D-amino-acid dehydrogenase [EC:1.4.99.1] | **Unknown function:** Enzymes of unknown specificity |
| PP_3597 | - | amino acid ABC transporter ATP-binding protein | K02028 polar amino acid transport system ATP-binding protein [EC:3.6.3.21] | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3598 | - | peptidase C26 | K07010 putative glutamine amidotransferase | **Protein fate:** Degradation of proteins, peptides, and glycopeptides |
| PP_3599 | - | 5-dehydro-4-deoxyglucarate dehydratase (EC:4.2.1.41) | K01707 5-dehydro-4-deoxyglucarate dehydratase [EC:4.2.1.41] | **Energy metabolism:** Other |
| PP_3600 | - | D-galactonate transporter | K03535 MFS transporter, ACS family, glucarate transporter | **Transport and binding proteins:** Carbohydrates, organic alcohols, and acids |
| PP_3601 | - | D-galactarate dehydratase, putative | K01708 galactarate dehydratase [EC:4.2.1.42] | **Energy metabolism:** Other |
| PP_3602 | - | ketoglutarate semialdehyde dehydrogenase | - | **Energy metabolism: (**Fermentation/other**)** |
| PP_3603 | - | GntR family transcriptional regulator | K05799 GntR family transcriptional regulator, transcriptional repressor for pyruvate dehydrogenase complex | **Regulatory functions:** DNA interactions |
| PP_3604 | - | hypothetical protein | - | - |
| PP_3605 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3606 | *qor-2* | quinone oxidoreductase | K00344 NADPH2:quinone reductase [EC:1.6.5.5] | **Energy metabolism:** Electron transport |
| PP_3607 | - | conserved hypothetical protein | K07112 | **Hypothetical proteins:** Conserved |
| PP_3608 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3609 | - | conserved hypothetical protein | K09936 hypothetical protein | **Hypothetical proteins:** Conserved |

| | | | | |
|---|---|---|---|---|
| PP_3610 | - | hypothetical protein | - | - |
| PP_3611 | - | hypothetical protein | - | - |
| PP_3612 | - | TonB-dependent siderophore receptor | K02014<br>iron complex outermembrane recepter protein | **Transport and binding proteins:** Cations and iron carrying compounds |
| PP_3613 | - | L-sorbosone dehydrogenase | - | **Energy metabolism:** Sugars |
| PP_3614 | - | hypothetical protein | - | **Unknown function:** General |
| PP_3615 | - | hypothetical protein | - | - |
| PP_3616 | - | hypothetical protein | - | - |
| PP_3617 | - | hypothetical protein | - | - |
| PP_3618 | - | hypothetical protein | - | - |
| PP_3619 | - | hypothetical protein | - | - |
| PP_3620 | - | conserved hypothetical protein | K09136<br>hypothetical protein | **Hypothetical proteins:** Conserved |
| PP_3621 | - | (2Fe-2S)-binding domain protein | K07302<br>isoquinoline 1-oxidoreductase, alpha subunit [EC:1.3.99.16] | **Energy metabolism:** Other |
| PP_3622 | - | isoquinoline 1-oxidoreductase, beta subunit, putative | K07303<br>isoquinoline 1-oxidoreductase, beta subunit [EC:1.3.99.16] | **Energy metabolism:** Other |
| PP_3623 | - | gluconate 2-dehydrogenase acceptor subunit | - | **Energy metabolism:** Electron transport |
| PP_3624 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3625 | - | amino acid transporter LysE | - | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3626 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3627 | - | hypothetical protein | - | - |
| PP_3628 | - | choline/carnitine/betaine transporter | - | **Transport and binding proteins:** Other |
| PP_3629 | - | 3-demethylubiquinone-9 3-methyltransferase | K00568<br>3-demethylubiquinone-9 3-methyltransferase [EC:2.1.1.- 2.1.1.64] | **Biosynthesis of cofactors, prosthetic groups, and carriers:** Menaquinone and ubiquinone |
| PP_3630 | - | porin, putative | - | **Transport and binding proteins:** Porins |
| PP_3631 | - | conserved hypothetical protein | K07184<br>SH3 domain protein | **Hypothetical proteins:** Conserved |

| | | | | |
|---|---|---|---|---|
| PP_3632 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3633 | *argC* | N-acetyl-gamma-glutamyl-phosphate reductase (EC:1.2.1.38) | K00145 N-acetyl-gamma-glutamyl-phosphate reductase [EC:1.2.1.38] | **Amino acid biosynthesis:** Glutamate family |
| PP_3634 | - | acetyltransferase | - | **Unknown function:** Enzymes of unknown specificity |
| PP_3635 | - | binding-protein-dependent transport systems inner membrane component | K02050 sulfonate/nitrate/taurine transport system permease protein | **Transport and binding proteins:** Anions |
| PP_3636 | - | sulfonate ABC transporter, periplasmic sulfonate-binding protein, putative | K02051 sulfonate/nitrate/taurine transport system substrate-binding protein | **Transport and binding proteins:** Anions |
| PP_3637 | - | sulfonate ABC transporter, ATP-binding protein, putative | K02049 sulfonate/nitrate/taurine transport system ATP-binding protein | **Transport and binding proteins:** Anions |
| PP_3638 | - | acyl-CoA dehydrogenase, putative | - | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3639 | - | alkylhydroperoxidase | - | **Unknown function:** General |
| PP_3640 | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3641 | - | cytosine/purine/uracil/thiamine/allantoin permease family protein | K03457 nucleobase:cation symporter-1, NCS1 family | **Transport and binding proteins:** Nucleosides, purines and pyrimidines |
| PP_3642 | - | hypothetical protein | - | - |
| PP_3643 | - | oxidoreductase, putative | - | **Unknown function:** Enzymes of unknown specificity |
| PP_3644 | - | luciferase family protein | - | **Unknown function:** General |
| PP_3645 | - | alpha/beta fold family hydrolase | - | **Unknown function:** Enzymes of unknown specificity |
| PP_3646 | - | aldehyde dehydrogenase family protein | - | **Energy metabolism:** Fermentation |
| PP_3647 | - | oxidoreductase, putative | - | **Unknown function:** Enzymes of unknown specificity |
| PP_3648 | - | carboxymuconolactone decarboxylase | K01607 4-carboxymuconolactone decarboxylase [EC:4.1.1.44] | **Energy metabolism:** Other |
| PP_3649 | - | GntR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3650 | - | NUDIX hydrolase | - | **DNA metabolism:** DNA replication, recombination, and repair |
| PP_3651 | - | GAF sensor hybrid histidine kinase | - | **Regulatory functions:** Small molecule interactions / **Signal transduction:** Two-component systems |

| | | | | |
|---|---|---|---|---|
| PP_3652 | - | amino acid transporter LysE | - | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3653 | - | amino acid transporter LysE | - | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3654 | - | AsnC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3655 | - | cytosine/purine/uracil/thiamine/allantoin permease family protein | - | **Transport and binding proteins:** Nucleosides, purines and pyrimidines |
| PP_3656 | - | aromatic compound-specific porin, putative | - | **Transport and binding proteins:** Porins |
| PP_3657 | - | nitrobenzoate reductase, putative | - | **Energy metabolism:** Other |
| PP_3658 | - | aromatic compound MFS transporter, putative | K05548 MFS transporter, AAHS family, benzoate transport protein | **Transport and binding proteins:** Other |
| PP_3659 | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3660 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3661 | - | hypothetical protein | - | **Cell envelope:** Other |
| PP_3662 | - | decarboxylase family protein | K06966 | **Unknown function:** Enzymes of unknown specificity |
| PP_3663 | - | GGDEF domain-containing protein | - | **Unknown function:** General |
| PP_3664 | *pssA* | phosphatidylserine synthase (EC:2.7.8.8) | K00998 phosphatidylserine synthase [EC:2.7.8.8] | **Fatty acid and phospholipid metabolism:** Biosynthesis |
| PP_3665 | - | AraC family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3666 | - | major facilitator transporter | K03762 MFS transporter, MHS family, proline/betaine transporter | **Transport and binding proteins:** Unknown substrate |
| PP_3667 | - | creatinase | - | **Protein fate:** Degradation of proteins, peptides, and glycopeptides |
| PP_3668 | - | catalase/peroxidase HPI | K03782 catalase/peroxidase [EC:1.11.1.6] | **Cellular processes:** Detoxification |
| PP_3669 | - | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3670 | - | Membrane putative protein | - | **Cell envelope:** Other |
| PP_3671 | - | aldo/keto reductase family oxidoreductase | K06222 2,5-diketo-D-gluconate reductase B [EC:1.1.1.274] | **Unknown function:** Enzymes of unknown specificity |
| PP_3672 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Domain |
| PP_3673 | - | transposase, OrfB, degenerate | - | **Mobile and extrachromosomal element functions:** Transposon functions / **Disrupted reading frame** |

| | | | | |
|---|---|---|---|---|
| PP_3674 | - | transposase, OrfA, authentic point mutation | - | **Mobile and extrachromosomal element functions:** Transposon functions |
| PP_3675 | - | hypothetical protein | - | - |
| PP_3676 | - | hypothetical protein | - | - |
| PP_3677 | - | hypothetical protein | - | - |
| PP_3678 | - | hypothetical protein | - | - |
| PP_3679 | - | hypothetical protein | - | - |
| PP_3680 | - | hypothetical protein | K07459 putative ATP-dependent endonuclease of the OLD family | - |
| PP_3681 | - | helicase, putative | - | **Unknown function:** General |
| PP_3682 | - | hypothetical protein | - | - |
| PP_3683 | - | hypothetical protein | - | - |
| PP_3684 | - | hypothetical protein | - | - |
| PP_3685 | - | hypothetical protein | - | - |
| PP_3686 | - | hypothetical protein | - | - |
| PP_3687 | - | ISPpu14, transposase, truncation | - | **Mobile and extrachromosomal element functions:** Transposon functions / **Disrupted reading frame** |
| PP_3688 | - | hypothetical protein | - | - |
| PP_3689 | - | serine/threonine protein phosphatase, putative | - | **Regulatory functions:** Protein interactions |
| PP_3690 | - | hypothetical protein | - | - |
| PP_3691 | - | DNA helicase-related protein | - | **Unknown function:** General |
| PP_3692 | - | hypothetical protein | - | - |
| PP_3693 | - | transcriptional regulator MvaT, P16 subunit, putative | - | **Regulatory functions:** DNA interactions |
| PP_3694 | - | hypothetical protein | - | - |
| PP_3695 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |

| | | | | |
|---|---|---|---|---|
| PP_3696 | - | hypothetical protein | - | - |
| PP_3697 | - | hypothetical protein | - | - |
| PP_3698 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3699 | - | hypothetical protein | - | - |
| PP_3700 | - | hypothetical protein | - | - |
| PP_3701 | - | hypothetical protein | - | - |
| PP_3702 | - | hypothetical protein | - | - |
| PP_3703 | - | hypothetical protein | - | - |
| PP_3704 | - | hypothetical protein | - | - |
| PP_3705 | - | hypothetical protein | - | - |
| PP_3706 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3707 | - | hypothetical protein | - | - |
| PP_3708 | - | hypothetical protein | - | - |
| PP_3709 | - | hypothetical protein | - | - |
| PP_3710 | - | hypothetical protein | - | - |
| PP_3711 | - | sensory box protein | - | **Regulatory functions:** Small molecule interactions |
| PP_3712 | - | conserved hypothetical protein | - | **Hypothetical proteins:** Conserved |
| PP_3713 | *catA* | catechol 1,2-dioxygenase | K03381 catechol 1,2-dioxygenase [EC:1.13.11.1] | **Energy metabolism:** Other |
| PP_3714 | *catC* | muconolactone isomerase | K03464 muconolactone D-isomerase [EC:5.3.3.4] | **Energy metabolism:** Other |
| PP_3715 | *catB* | muconate and chloromuconate cycloisomerase | K01856 muconate cycloisomerase [EC:5.5.1.1] | **Energy metabolism:** Other |
| PP_3716 | *catR* | LysR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3717 | - | LuxR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3718 | - | hypothetical protein (jcvi: aminotrasnferase, classIII) | - | **Unknown function:** Enzymes of unknown specificity |

| | | | | |
|---|---|---|---|---|
| PP_3719 | - | periplasmic polyamine-binding protein, putative | - | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3720 | - | ribosyldihydronicotinamide dehydrogenase (quinone) | K00355 NAD(P)H dehydrogenase (quinone) [EC:1.6.5.2] | **Energy metabolism:** Electron transport |
| PP_3721 | *aspC* | aspartate aminotransferase | K12252 arginine:pyruvate transaminase [EC:2.6.1.84] | **Amino acid biosynthesis:** Aspartate family |
| PP_3722 | *alr* | alanine racemase (EC:5.1.1.1) | K01775 alanine racemase [EC:5.1.1.1] | **Cell envelope:** Biosynthesis and degradation of murein sacculus and peptidoglycan |
| PP_3723 | - | acetolactate synthase, large subunit, putative, authentic frameshift | - | **Amino acid biosynthesis:** Pyruvate family |
| PP_3724 | - | acyl-CoA synthetase, putative | - | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3725 | - | acyl-CoA dehydrogenase domain-containing protein | K00257 [EC:1.3.99.-] | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3726 | - | enoyl-CoA hydratase/isomerase | K01692 enoyl-CoA hydratase [EC:4.2.1.17] | **Fatty acid and phospholipid metabolism: D**egradation |
| PP_3727 | - | amino acid transporter | K02205 arginine/ornithine permease | **Transport and binding proteins:** Amino acids, peptides and amines |
| PP_3728 | - | multi-sensor hybrid histidine kinase | - | **Regulatory functions:** Small molecule interactions / **Signal transduction:** Two-component systems |
| PP_3729 | - | periplasmic amino acid-binding protein-related protein | K02030 polar amino acid transport system substrate-binding protein | **Unknown function:** General |
| PP_3730 | - | winged helix family two component transcriptional regulator | - | **Regulatory functions:** DNA interactions / **Signal transduction:** Two-component systems |
| PP_3731 | - | TetR family transcriptional regulator | - | **Regulatory functions:** DNA interactions |
| PP_3732 | - | enoyl-CoA hydratase/isomerase | K01692 enoyl-CoA hydratase [EC:4.2.1.17] | **Fatty acid and phospholipid metabolism:** Degradation |
| PP_3733 | - | ABC transporter lipoprotein, putative | - | **Transport and binding proteins:** Unknown substrate |

For Table 8-8 and Table 8-9 different colors were employed in order to simplify the visualization of the cellular roles of the different encoded proteins. The colors were listed below in alphabetical order:

Amino acid Biosynthesis

Biosynthesis of cofactors, prosthetic groups and carriers

Cell envelope

Cellular processes

Central intermediary metabolism

Disrupted reading frame

DNA metabolism
Energy metabolism
Fatty acid and phospholipid metabolism
Mobile and extrachromosomal element functions
Protein fate
Regulatory functions
Signal transduction
Transcription
Transport and binding proteins

When an encoded protein participated to two different cellular roles, the cell was left empty in the Tables. In the same way when a gene encoded a hypothetical or conserved hypothetical protein as well as a protein with unknown function, the cells were left in white color.

## 8.5 Bioscreen experiments

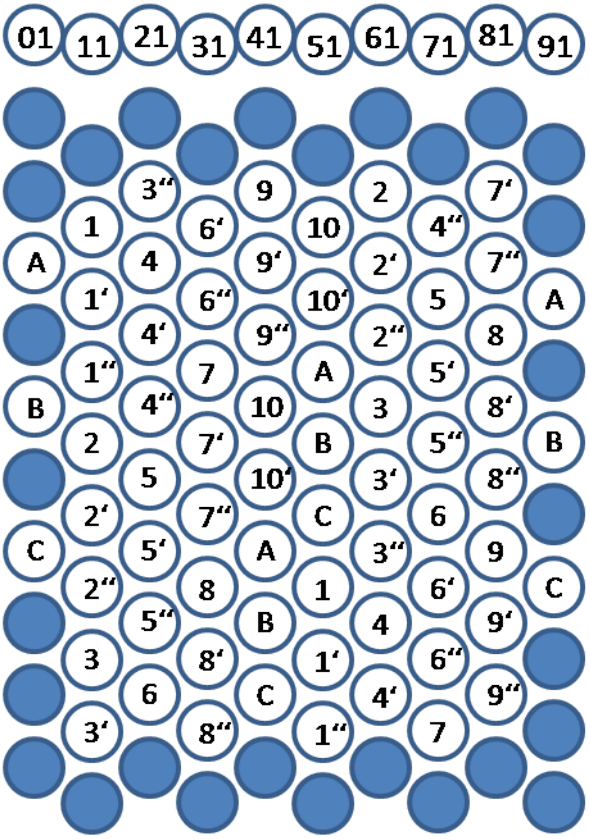



**Figure 8-1: Organization of one honeycomb plate used for establishing the growth curves of *P. putida* wild-type and mutant strains performed with the Bioscreen**
The circles colored in blue represent wells filled up with water. The letters A, B and C represent the control samples, the wells are filled up with the M9 medium and different carbon sources. The numbers correspond to different strains in the different conditions.

Duplicates of a single strain were inoculated on a same plate for each carbon and amino acid source. The numbers on Figure 8-1 referred to Table 8-10. The 100-well honeycomb plate drawn on the Figure 8-1 is shown as an example for the inoculation of *P. putida* wild-type and mutant strains. In this case precisely the $\Delta_x$-407 mutant series is illustrated. The circles colored in blue represented the wells filled up with distillated water. The blue circles with white background were filled up with the different media and the numbers represented the strains, as described in Table 8-10. The letters A, B and C correspond to the blanks done with the different carbon sources.

**Table 8-10: List of the carbon and amino acid sources used for the growth in the Bioscreen of the *P. putida* TEC1 wild-type strain and *P. putida* $\Delta_x$-407 mutants series**

| Strains / Carbon/ amino acid sources | | *P. putida* $\Delta_1$ 407 | *P. putida* $\Delta_2$-407 col.1 | *P. putida* $\Delta_2$-407 col.3 | *P. putida* TEC1 |
|---|---|---|---|---|---|
| **Citrate** | Ø | 1 | 4 | 7 | 10 |
| | L-Arg | 1′ | 4′ | 7′ | 10′ |
| | L-Val + L-Ile | 1′′ | 4′′ | 7′′ | 10′′ |
| **Glucose** | Ø | 2 | 5 | 8 | 11 |
| | L-Arg | 2′ | 5′ | 8′ | 11′ |
| | L-Val + L-Ile | 2′′ | 5′′ | 8′′ | 11′′ |
| **Succinate** | Ø | 3 | 6 | 9 | 12 |
| | L-Arg | 3′ | 6′ | 9′ | 12′ |
| | L-Val + L-Ile | 3′′ | 6′′ | 9′′ | 12′′ |

In Table 8-10, numbers are associated with the different *P. putida* strains (wild-type and mutants) for the different carbon and amino acid sources. These numbers correspond to the ones drawn on the plate drawn in Figure 8-1. In total five Carbon sources were used: citrate, glucose, succinate, benzoate and glycerol and a sixth condition corresponded to the M9 minimal medium without and carbon source supplement (represented as Ø in the tables). Only the three first C sources are represented as an example. For each C source three different conditions were applied corresponding to the addition or not of certain amino acids: L-arginine, L-valine (with a trace of L-isoleucine) and finally no amino acid (represented by the symbol: Ø).

| | |
|---|---|
| Name, Vorname: | **LEPRINCE, Audrey** |
| Geburts(datum/ort): | 7.Oktober 1983 in Créteil (F) |
| Staatsangehörigkeit: | französisch |

## Ausbildung

| | |
|---|---|
| 07/2001 | Naturwissenschaftliches Abitur in Physik und Vordiplom DEUG (im Bereich der Biologie) |
| 09/2001-06/2003 | Vorbereitungsklasse für die Aufnahmeprüfung für „Grandes Ecoles“ am Gymnasium „Marcelin Berthelot“ in Saint-Maur (F)<br>Hauptfächer: Biologie, Mathematik, Chemie, Physik |
| 09/2003-06/2005 | Studium der Ingenieurwissenschaften an der Ingenieurhochschule ENSIA (Ecole Nationale Supérieure des Industries AgroAlimentaires) |
| 10/2005-03/2006 | Studium der Lebensmitteltechnologie an der TU München (D) |
| 01/2007-11/2010 | Doktorandin am Helmholtz Zentrum für Infektionsforschung, Braunschweig (D) (Molekulare Biotechnologie, Mikrobiologie, Genetik) |

## Praktika

| | |
|---|---|
| 03/2006-09/2006 | Diplomarbeit am HZI in der Gruppe für System- und Synthetische Biologie (RNA Gel Elektrophorese, RT-PCR…) |
| 09/2004-12/2004 | Praktikum in der Zuckerraffinerie SDHF in Attin (F) |
| 07-08/2004 | Praktikum bei Ravifruit (F), Unternehmen für die Weiterverarbeitung den Früchten (Logistik) |
| 08/2003 | Praktikum in der Bank Crédit du Nord in Paris (F) |

### EDV-Kenntnisse

| | |
|---|---|
| Betriebssysteme: | Windows, UNIX |
| Software: | Word, Works, Excel end PowerPoint |

## Sprachkenntnisse

| | |
|---|---|
| Französisch: | Muttersprache |
| Deutsch: | sehr gut in Wort und Schrift |
| Englisch: | sehr gut in Wort und Schrift |

## Publikationen

Suárez Diez, M., C. M. C. Lam, A. Leprince & Martins dos Santos, V. A. P. “(Re-) construction, characterization and modeling of systems for synthetic biology.” Biotechnology Journal 4(10): 1382-1391 (2009)

Leprince, A., D. Janus, V. de Lorenzo & Martins dos Santos, V. A. P. “Streamlining of a *Pseudomonas putida* genome using a combinatorial method based on mini-transposon insertion and the Flp-*FRT* recombination system.” Invited book chapter in Synthetic Gene Networks